EUROPEAN PARLIAMENT

Mass Surveillance

Part 1 - Risks, Opportunities and Mitigation Strategies

ANNEX

Science and Technology Options Assessment

EPRS | European Parliamentary Research Service

Scientific Foresight (STOA) Unit

PE 527.409

Mass Surveillance

What are the risks for the citizens and the opportunities for the European Information Society? What are the possible mitigation strategies?

Part 1 - Risks and opportunities raised by the current generation of network services and applications

Annex

IP/G/STOA/FWC-2013-1/LOT 9/C5/SC1

January 2015

PE 527.409

The STOA project 'Mass Surveillance - Part 1: Risks and opportunities raised by the current generation of network services and applications' was carried out by TECNALIA Research and Investigation.

AUTHORS

Arkaitz Gamino Garcia
Concepción Cortes Velasco
Eider Iturbe Zamalloa
Erkuden Rios Velasco
Iñaki Eguía Elejabarrieta
Javier Herrera Lotero
José Javier Larrañeta Ibañez
Stefan Schuster (Editor)

STOA RESEARCH ADMINISTRATOR

Peter Ide-Kostic
Scientific Foresight Unit
Directorate for Impact Assessment and European Added Value
Directorate-General for Parliamentary Research Services
European Parliament, Rue Wiertz 60, B-1047 Brussels
E-mail: peter.ide-kostic@europarl.europa.eu

LINGUISTIC VERSION

Original: EN

ABOUT THE PUBLISHER

To contact STOA or to subscribe to its newsletter please write to: STOA@ep.europa.eu
This document is available on the Internet at: http://www.ep.europa.eu/stoa/

Manuscript completed in January 2015
Brussels, © European Union, 2015

DISCLAIMER

The content of this document is the sole responsibility of the author and any opinions expressed therein do not necessarily represent the official position of the European Parliament. It is addressed to the Members and staff of the EP for their parliamentary work. Reproduction and translation for non-commercial purposes are authorised, provided the source is acknowledged and the European Parliament is given prior notice and sent a copy.

PE 527.409
ISBN: 978-92-823-5537-4
DOI: 10.2861/566206
CAT: QA-05-14-156-EN-N

Abstract

This document contains the Annex to the Study on Mass Surveillance, commissioned by STOA of the European Parliament. This Annex contains detailed information and references that were used as the basis for the elaboration of the Study.
The motivation for providing this Annex in conjunction with the Study is to provide the reader with a means to delve deeper into relevant information concerning the questions posed in the tender of the Study.
In order to facilitate the reader an appropriate orientation and guidance for accessing this information, this Annex is structured according to the themes and questions specified in the original tender. Specific information or crosslinks to information that is relevant for more than one question are provided for each of the thirty five questions documented in the tender.

CONTENTS

LIST OF ABBREVIATIONS 1
LIST OF TABLES 3
LIST OF FIGURES 4
1 THEME 1 - CURRENT PRACTICES OF INTERCEPTION AND ANALYSIS OF END-USER META-DATA 5
1.1 Question 1 5
1.2 Question 2 10
1.3 Question 3 13
1.4 Question 4 17
1.5 Question 5 19
1.6 Question 6 20
2 THEME 2 - CRYPTOGRAPHY RELIABILITY IN A "POST SNOWDEN" WORLD 33
2.1 Question 7 33
2.2 Question 8 38
2.3 Question 9 41
2.4 Question 10 42
2.5 Question 11 44
2.6 Question 12 45
2.7 Question 13 46
2.8 Question 14 49
2.9 Question 15 51
2.10 Question 16 53
2.11 Question 17 53
2.12 Question 18 54
3 THEME 3 - EFFICIENCY OF COMMERCIAL PRODUCTS FOR DOING MASS SURVEILLANCE 56
3.1 Question 19 56
3.2 Question 20 68
3.3 Question 21 75
3.4 Question 22 79
3.5 Question 23 80
3.6 Question 24 83
3.7 Question 25 86

3.8 Question 26 87
3.9 Question 27 88
3.10 Question 28 90
3.11 Question 29 90
4 THEME 4 - TECHNICAL CREDIBILITY OF NATIONAL SECURITY AGENCIES HACKING CAPABILITIES 92
4.1 Question 30 92
4.2 Question 31 93
4.3 Question 32 96
4.4 Question 33 99
4.5 Question 34 102
4.6 Question 35 104

LIST OF ABBREVIATIONS

ACLU	American Civil Liberties Union
ADSL	Asymmetric Digital Subscriber Line
AES	Advanced Encryption Standard
APT	Advanced Persistent Threat
BIOS	Basic Input/Output System
BND	Bundesnachrichtendienst (German foreign intelligence agency)
CDD	Center for Digital Democracy
CGHQ	Government Communications Headquarters (UK)
DAA	Digital Advertising Alliance
DDoS	Distributed Denial of Service
DNT	Digital Network Technologies
DOD	(US) Department of Defense
DPI	Deep Packet Inspection
DRIPA	Data Retention and Investigatory Powers Act
EB	Exabytes
EFF	Electronic Frontier Foundation
ETSI	European Telecommunications Standard Institute
FISC/FISA	Foreign Intelligence Surveillance Court/Act
FPGA	Field-programmable Gate Array
FTC	Federal Trade Commission
FVEY	Five Eyes Intelligence Alliance (USA, Canada, UK, Australia, New Zealand)
GERD	GastroEsophageal Reflux Disease
GPRS	General Packet Radio Service
GPS	Global Positioning System
HIx	Handover Interface x
HTTP	Hypertext Transfer Protocol
IC	Integrated Circuit
IC	Intelligence Community
IDF	Israeli Defence Forces
IETF	Internet Engineering Task Force
IMAP	Internet Message Access Protocol
IMEI	International Mobile station Equipment Identity
IMSI	International Mobile Subscriber Identity
IP	Internet protocol
IPv4/6	Internet Protocol version 4/6
IRI	Intercept Related Information
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
ITOM	IT Operation and Management
LAN	Local Area Network
MAC	Media Access Control
MitM	Man in the Middle (attack)
NI	Network Inspection
NOFORN	NO FOReign National access allowed (security classification)
NSA	National Security Agency (of the USA)
OSI	Open Systems Interconnection
PDP	Packet Data Protocol
PDU	Protocol Data Unit

PFS	Perfect Forward Secrecy
PGP	Pretty Good Privacy (open source encryption tool)
POP3	Post Office Protocol (version 3)
PSTN	Public Switched Telephone Network
REL (TO X)	Release to country X allowed (security classification)
RFC	Request for Comment
RIPA	Regulation of Investigatory Powers Act
RNG	Random Number Generator
SCA	Stored Communications Act
SI	Special Intelligence
SIGAD	Signals Intelligence Activity Designator
SIGINT	Signals Intelligence
SIM	Subscriber Identity Module
SMTP	Simple Mail Transfer Protocol
SSL	Secure Socket Layer
TAO	Tailored Access Operations (a division of NSA)
TOR	The Onion Ring Network
TPC	Transmission Control Protocol
TS	Top Secret (security classification)
UDP	User Datagram Protocol
USB	Universal Serial Bus
voIP	Voice over Internet Protocol
VPN	Virtual Private Network
xDSL	different variants of Digital Subscriber Lines

LIST OF TABLES

Table 1: Data types that can be intercepted depending on the OSI layer 7
Table 2: Relation between password length and effort for cracking it (source theWire) 37
Table 3: Encryption practices of major internet companies (source EFF) 48
Table 4: Suggested cryptoperiods for key types (source NIST[251]) 52
Table 5: Catalogues of surveillance vendors and products 63
Table 6: Feature-set of different FinFisher components (source Gama Group) 71

LIST OF FIGURES

Figure 1: IP metadata (labelled as HI2) associated to each OSI layer 7

Figure 2: IAB proposal of solution classes to replace traditional cookies 16

Figure 3: Schematic diagram of NSA's interception nodes 22

Figure 4: Duplicated 'got fail' statement in Apple's SSL verification code 34

Figure 5: CANDYGRAM Operational Concept 42

Figure 6: Browser support for PFS (as of June 2013 - source Netcraft) 50

Figure 7: Timeline of surveillance deployment in Tunisia 57

Figure 8: Lawful Interception, Intelligence Gathering, Telecom Networks and ISS Vendors (source Telestrategies) 59

Figure 9: Lawful Interception and Intelligence Gathering Vendors (source Telestrategies) 60

Figure 10: Global Proliferation of FinFisher surveillance suite 61

Figure 11: Blue Coat gobal deployment 62

Figure 12: Surveillance companies and technologies in the UK (source Wikileaks Spyfiles-map) 63

Figure 13: Components of FinFisher product portfolio (source Gamma Group) 69

Figure 14: Steps for using FinUSB (source GammaGroup) 72

Figure 15: ELAMAN's POSEIDON Internet Monitoring Centre connected with a tap to the network (source Privacy International[324]) 74

Figure 16: AMESYS' EAGLE System (source Privacy International) 74

Figure 17: AMESYS' EAGLE System's Man-Machine Interface (source Privacy International[325]) 75

Figure 18: Screenshots of email, chat and voip sessions reconstructed with ELAMAN POSEIDON (Source Privacy international[324]) 75

Figure 19: Screenshot of BOUNDLESS INFORMANT 77

Figure 20: Communication monitoring capabilities in Germany 77

Figure 21: Monitoring capacity of telephone metadata in Spain and Italy 78

Figure 22: Current volume and limits of FASCIAs telephony metadata storage capacity 78

Figure 23: 7 means for defeating encryption (source TeleStrategies) 81

Figure 24: Verint marketing information 84

Figure 25: Nice marketing information 85

Figure 26: Eagle product brochure differentiating between lawful and massive interception 85

Figure 27: ClearTrail corporate information 86

Figure 28: Extract of slide #11 of Treasuremap presentation leaked by Snowden 94

Figure 29: Leaked slide of a presentation to the 2011 annual conference of the Five Eyes alliance on NSA's SIGINT gathering strategy 100

Figure 30: Leaked slide on GCHQs strategy on breaking encryption technology 103

1 THEME 1 - CURRENT PRACTICES OF INTERCEPTION AND ANALYSIS OF END-USER META-DATA

1.1 Question 1

What type of meta-data can be gathered for mass surveillance purposes nationally and internationally over what type of infrastructures (ie, private terrestrial and satellite telecom operators, the public Internet, wireless mobile service operators, etc...)?
How does the interception practically and technically work?

Metadata is often called *"data about data"* and *"is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource"*[1].

The definition of metadata can be approached from a technical perspective or a legal perspective. The comprehensive analysis by M. Bellovin of Columbia University[2] shows that metadata cannot be defined solely in technical terms, but technical considerations must play an important role in creating such definition.

A 2008 report by US National Research Council of the National Academies on privacy catalogued the forms of metadata and data created about individuals:

"[…] ***financial transactions, medical records, travel, communications, legal proceedings, consumer preferences, Web searches, and, increasingly, behavioral and biological information****. This is the essence of the information age – it provides us with convenience, choice, efficiency, knowledge, and entertainment; it supports education, health care, safety, and scientific discovery. Everyone leaves personal digital tracks in these systems whenever he or she makes a purchase, takes a trip, uses a bank account, makes a phone call, walks past a security camera, obtains a prescription, sends or receives a package, files income tax forms, applies for a loan, e-mails a friend, sends a fax, rents a video, or engages in just about any other activity"*[3]

In the context of communications, the information resource can refer to either the content (communicated message) or the communication itself. Therefore, **there are two types of metadata** which are frequently confused in the literature: **metadata that provides data on the content** (e.g. read/write/modify attributes of the file, author of the document, GPS location of the photo scene, etc.), **and metadata of the communication** (e.g. sender, receiver, communication duration, communication starting date and time, communication channel, communication protocol used, etc.).

When it comes to mass surveillance, it is the communication metadata that is being subject to discussion about whether it is more revealing of private information than the content itself. In fact, **from a legal perspective, the communication metadata is the only existing metadata,** as metadata of the content is considered as part of the content, which travels end-to-end embedded in the content.

Communication meta-data, both with respect to internet or telephone communication (terrestrial, satellite, or wireless), is routinely gathered by the respective telecom providers and ISPs as part of their business operations. The Directive 2006/24/EC[4] of the European Parliament established a mandatory retention period of this meta-data for a minimum of 6 months and a maximum of 2 years by all communication providers in European Member Countries. This Directive has been declared

[1] http://www.niso.org/publications/press/UnderstandingMetadata.pdf
[2] https://www.cs.columbia.edu/~smb/papers/PCLOB-statement.pdf
[3] http://epic.org/misc/nrc_rept_100708.pdf
[4] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:EN:PDF

invalid by the Court of Justice of the European Union on 8 April 2014[5], arguing that *by requiring the retention of those data and by allowing the competent national authorities to access those data,* ***the directive interferes in a particularly serious manner with the fundamental rights to respect for private life and to the protection of personal data***, while recognizing that *the retention of data for the purpose of their possible transmission to the competent national authorities genuinely satisfies an objective of general interest, namely the fight against serious crime and, ultimately, public security.*

After the court's decision of invalidating the Directive 2006/24/EC**, the UK has approved** the Data Retention and Investigatory Powers Act 2014 (DRIP Act), which establishes **that providers can be required to retain meta-data for up to 12 months** and that this data may be used as evidence by law enforcement. The DRIP Act also settles that any communication service provider, regardless of its provenance, must comply with lawful requests under the UK Regulation of Investigatory Powers Act (RIPA) if providing services to UK customers.

Australia is currently discussing whether to introduce **data retention legislation** that would require providers to store meta-data (i.e. IPs) for up to 2 years.

There is no legal obligation for data retention by communication providers in the USA. However, **US law enforcement can obtain access to meta-data stored by providers under the Stored Communications Act**[6] (SCA), which also establishes mandatory data storage for up to 180 days upon government request.

In the context of intelligence services, metadata is data that "*describes content, events, or networks associated with SIGINT targets*" as defined in a 2009 draft report prepared by the NSA's inspector general[7]. This report also identifies **two different types of communication metadata: Telephony metadata and Internet metadata** (also called Internet Protocol (IP) metadata). For example, for an email message, the metadata would include the sender and recipient email addresses. It does not contain the subject line or the text of the email which are considered to be content. Likewise, for a telephone conversation metadata would at least include the called number and the calling number, as well as the duration of the call.

As described in a declassified order ruled by the FISA Court which required collecting telephony metadata of a company, the **telephony metadata includes "*comprehensive communications routing information (e.g. originating and terminating telephone number, International Mobile Subscriber Identity (IMSI) number, International Mobile station Equipment Identity (IMEI) number, etc.), trunk identifier, telephone calling card numbers, and time and duration of the call.*** *Telephony metadata does not include the content of the communication, or the name, address, or financial information of a subscriber or customer"*[8]. The metadata information, also classed as transactional information, can also include the cell site location data, i.e. the nearest cell tower a phone was connected to[9].

IP metadata can be collected for services that run over IP protocol (e.g. VoIP, email, HTTP-based services such as facebook, twitter, searches over internet, etc.) or that IP runs on top of (e.g. GPRS, Ethernet, etc.)[10].

In contrast to telephony metadata, **multiple parties are involved in IP metadata interception** since the transportation of IP-based services over the Internet is done by Internet access providers on each of the communications, networks operators, and providers of the services themselves[11]. Moreover, the separation of the IP communication in several communication layers (Open Systems Interconnection

[5] http://curia.europa.eu/jcms/upload/docs/application/pdf/2014-04/cp140054en.pdf
[6] http://www.gpo.gov/fdsys/pkg/STATUTE-100/pdf/STATUTE-100-Pg1848.pdf
[7] https://www.aclu.org/files/natsec/nsa/20130816/NSA%20IG%20Report.pdf
[8] http://www.dni.gov/files/documents/0708/BR%2009-09%20Primary%20Order.pdf
[9] http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order
[10] https://tools.ietf.org/html/rfc3924
[11] AQSACOM, Lawful interception for IP networks, White Paper. March 2010.

(OSI) layers[12]) makes it difficult to establish and identify which IP metadata can be gathered. ETSI published a set of specifications in order to support IP lawful interception for telecommunication service providers and network providers[13]. **The lawful interception of metadata is a targeted surveillance required by Law Enforcement Authorities and is not considered as mass surveillance.** The ETSI model can still be used for the identification of the IP metadata in each of the communication layers as the communication protocol is the same.

The set of standards defined by ETSI includes the definition of a **high level architecture for lawful interception in IP networks** which describes a minimum set of common internal network functions and interfaces[14]. Among the defined interfaces is the Handover Interface 2 (HI2), which is an interface port to transport all Intercept Related Information (IRI). ETSI defines IRI as the information or data associated with the communication services of the target identity apparent to the network, i.e. the metadata associated to the IP communication. They have also defined Handover Interface 3 (HI3) which is the content of the communication.

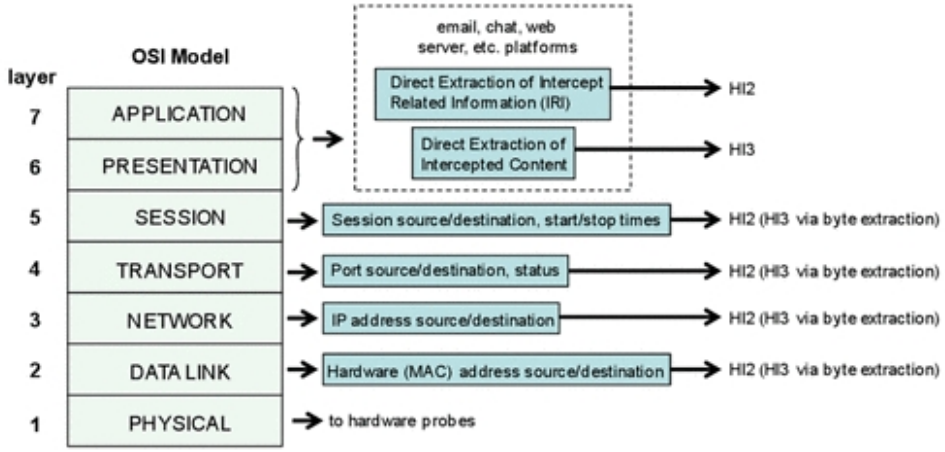


Figure 1: IP metadata (labelled as HI2) associated to each OSI layer

ETSI also identifies the data type that can be intercepted depending on the OSI layer and which party can be involved in it (access provider, network provider or service provider)[15].

Component	OSI Layer(s)	Format of intercepted data
Access provider	1 (Physical)	Physical PDUs
	2 (Data link)	Data link PDUs
	3 (Network)	(IP) Datagrams
Network connectivity	3 (Network)	(IP) Datagrams
Service Provider	5/7 (Application)	Application layer transactions

Table 1: Data types that can be intercepted depending on the OSI layer

[12] Recommendation, I. T. U. T. X. 200 (1994)| ISO/IEC 7498-1: 1994. Information technology–Open Systems Interconnection–Basic Reference Model: The basic model

[13] http://www.etsi.org/technologies-clusters/technologies/security/lawful-interception

[14] ETSI TR 102 528: "Lawful Interception (LI) Interception domain Architecture for IP networks".

[15] ETSI TS 102 232-1 V3.7.1. July 2014

Depending on the IP-based service and the type of access technology utilized, **the metadata that can be gathered** is different, as it **depends on the stack of protocols on which the communication service works**.

The type of access technology (e.g. GPRS, Ethernet or xDSL access) utilized in the communication will determine the metadata (that can be gathered) related to the OSI layer 2.

- Circuit switched network (e.g. xDSL) metadata:
    - PSTN/ISDN number for dial-up
    - MSISDN (Mobile Station Integrated Services Digital Network)
    - IMSI (International Mobile Subscriber Identity)
    - IMEI (International Mobile station Equipment Identity)
    - Event type (Establishment, Answer, Supplementary service, Handover, Release, SMS, Location update, Subscriber controlled input)
    - Event date
    - Event time
    - Dialled number (Dialled number before digit modification)
    - Connected number
    - Forwarded to number
    - Cell ID
    - Location area code
    - Serving system
- Packet switched network (e.g. GPRS, Ethernet) related metadata:
    - MAC address
    - MSISDN (Mobile Station Integrated Services Digital Network)
    - IMSI
    - IMEI
    - PDP (Packet Data Protocol) address(es). In case of IPv4v6 two addresses may be carried
    - Event type: PDP Context Activation, PDP Context Deactivation, GPRS Attach, etc.
    - Event date
    - Event time
    - Access point name[16]
    - PDP type[17]
    - Cell Global ID
    - Routing Area information (Routing Area in a GPRS-PLMN)
    - NSAPI (Network layer Service Access Point Identifier)

From the OSI layers 3 (network) and 4 (transport) of an IP communication the metadata that can be gathered is the following[18]:

- The type of Internet access (e.g. dial-up, ADSL, cable modem, LAN access)
- IP version (IPv4, IPv6)

[16] 3GPP TS 23.060

[17] 3GPP TS 09.60, 3GPP TS 29.060, 3GPP TS 04.08, 3GPP TS 24.008, 3GPP TS 09.02, 3GPP TS 29.002

[18] ETSI TS 102 232-3 v3.3.1. Lawful Interception (LI); Handover Interface and Service-Specific Details (SSD) for IP delivery; Part 3: Service-specific details for internet access services. October 2013.

- Source IP Address
- Destination IP Address
- Identification of the device used to communicate (e.g. computer name, DHCP Relay Agent Information)
- The date and time of the start of the session
- The date and time of a predicted session ending (e.g. lease expiration)
- The date and time of the end of the session (or lease)
- The reason for the session to end (e.g. logoff, connection loss, time out, lease expiration)
- The number of octets the target sent during the session
- The number of octets the target received during the session
- Source port
- Destination port
- Session status
- The MAC address of the target CPE for layer 2 access or the target PSTN/ISDN number for dial-up

This metadata is not technological or application dependent as it is the core of the IP communication and it is based on the stack of IP and TCP/UDP protocols. The type of the application service (e.g. email, HTTP-based web, VoIP) will determine the metadata (that can be gathered) related to OSI layers from 5 to 7.

For email services, the following IP metadata can be collected[19],[20]:

- server IP
- client IP
- server port
- client port
- email Protocol ID (SMTP, POP3, IMAP, Webmail)
- email sender address
- email recipient list
- server octets sent
- client octets sent
- Status

Internet metadata also includes:

1. information about IP address of the device from which an email or other electronic communication was sent (such as instant messaging), IP address destination and the IP address of routers and servers on the Internet that handled the communication during the transmission
2. the exchange of an IP address and email address that occurs when a user logs into a web-based email service
3. for certain logins to web-based email accounts, inbox metadata that is transmitted to the user upon accessing the account[21].

[19] ETSI TS 102 232-2 V3.7.1, Lawful Interception (LI); Handover Interface and Service-Specific Details (SSD) for IP delivery; Part 2: Service-specific details for messaging services. February 2014.
[20] http://www.thewire.com/technology/2013/06/email-metadata-nsa/66657/

An easy to understand explanation of metadata classification in internet communications has been published by "The Guardian" and differentiates among metadata of the most used types of services: email, phone, camera, facebook, twitter, search and web browser[22].

1.2 Question 2

> ***To which extent do organisations doing mass surveillance cooperate with famous third party data trade organisations such as Palantir but also on-line advertisement agencies such as Rubicon, Criteo, PubMatic, Appnexus in order to exploit the meta-data collected by these companies?***
> ***What form of cooperation takes place in practice?***

On-line advertisement agencies are becoming increasingly suspect of applying too aggressive or "immersive" practices that could be considered on the edge of mass surveillance of (potential) consumers. The ever evolving powerful techniques they use for advertising, such as data gathering, data analysis, consumer profiling (on demographics, geographic, emotional, contextual/semantic, etc. aspects), consumer behavioural tracking, and interactive advertisement targeting (in web banners, pop-ups and floating advertisements, etc.) **have given the advertising companies the ability to tailor the publicity to individuals at the cost of their privacy disclosure**.

"*Advertising strategies, campaigns, and distribution are increasingly based on predictive algorithms, spreadsheets, and math…. Every Web page's individual views, every word typed in a search query box (also known as the 'database of consumer intentions'), every video download, and even every word in an e-mail may create one more data point that a marketer can leverage and use to more precisely target the audience….*"[23]

As explained in a report by online marketer AppNexus, "[…] *Internet ad exchanges* […] *are basically markets for eyeballs on the Web.* ***Advertisers bid against each other in real time for the ability to direct a message at a single Web surfer****. The trades take 50 milliseconds to complete.*"[24]

"*We can build* [consumer] *profiles from any building blocks",* says Meir Zohar, chief executive of eXelate. "*Age, gender, purchase intent, interests, parents, bargain shoppers-you can assemble anything.*" eXelate "***gathers online consumer data through deals with hundreds of Web sites****. The firm determines a consumer's age, sex, ethnicity, marital status and profession by scouring Web-site registration data. It pinpoints, for example, which consumers are in the market to buy a car or are fitness buffs, based on their Internet searches and the sites they frequent.* ***It gathers and stores the information using tracking cookies, or small strings of data that are placed on the hard drive of a consumer's computer*** *when that consumer visits a participating site. Advertisers, in turn, purchase cookie data from eXelate and use it to buy targeted online ads.*"[25]

Health and finance sensitive information is also under the radar of advertisers' inference techniques, e.g. "*Google's DoubleClick Ad Exchange permits the targeting of a wide range of health and financial behaviors.*

---

[21] http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/6

[22] http://www.theguardian.com/technology/interactive/2013/jun/12/what-is-metadata-nsa-surveillance#meta=1111111

[23] Edward Landry, Carolyn Ude, and Christoper Vollmer, "HD Marketing 2010: Sharpening the Conversation," Booz/Allen/Hamilton, ANA, IAB, AAAA, 2008, http://www.boozallen.com/media/file/HD_Marketing_2010.pdf.

[24] Garrett Sloane, "amNY Special Report: New York City's 10 Hottest Tech Startups," amNewYork, 25 Jan. 2010, http://www.amny.com/urbanite-1.812039/amny-special-report-new-york-city-s-10-hottest-tech-startups-1.1724369

[25] H.R. 5777, The ''Best Practices Act,'' and H.R. lll, A discussion draft to require notice to and consent of an individual prior to the collection and disclosure of certain personal information relating to that individual, http://www.gpo.gov/fdsys/pkg/CHRG-111hhrg78124/pdf/CHRG-111hhrg78124.pdf

These include arthritis, diabetes, GERD and digestive disorders, migraines, sleep disorders, pain management, credit cards, loans and insurance."[25]

This situation led in April 2010 to a number of privacy groups, the Center for Digital Democracy, U.S. PIRG, and the World Privacy Forum to file a formal complaint[26] with the U.S. Federal Trade Commission (FTC), urging the FTC to investigate data collection "Wild West" involving real-time advertising auctions and data exchanges. The complaint cites companies such as Google, Yahoo, PubMatic[27], TARGUSinfo (now Neustar)[28], MediaMath[29], eXelate[30], Rubicon Project[31], AppNexus[32], and Rocket Fuel[33].

Since then, criticisms of the way online advertising companies[34] and internet service providers like Google[35] do perform online behavioural tracking and exchange of users' private information are continuously appearing. One of the most active agents in this field is the Center for Digital Democracy (CDD) which watches over the privacy of internet consumers. CDD has urged the FTC to take action on "*the use of geo-fencing, "geobehavioral targeting," "geo-cookies" and the role of location analytics, especially when integrated into broader data gathering* […]"[36].

CDD has raised their concern with respect to the network providers' involvement in data collection and targeted online marketing practices when deep packet inspection is used. "*Given actions by the FCC, consumers must rely on a handful of cable and telephone networks for their broadband service. Deep packet inspection (DPI) technologies enable these network providers to track their subscribers' actions online (data that can then be merged with extensive customer information files).* ***When the power of online ad profiling and targeting technologies are combined with the microscopic tracking and analysis capabilities of DPI, consumer privacy is further threatened***"[37].

Even though the FTC promoted in 2010 the "Do Not Track"[38] mechanism to let the users notify the websites about their preferences for not being tracked, the majority of websites has not yet committed to honouring this mechanism.

The advertisement market pleads for self-regulation and some advertisers in the Digital Advertising Alliance (DAA) are engaged in the DAA's Self-Regulatory Program for Online Behavioral Advertising[39] and include in their advertisements an advertising option icon that shows the consumer an Opt-Out Page[40]. The complete list of participating companies can be found online[41].

The Whitepaper on web tracking and privacy, released by the International Working Group on Data Protection in Telecommunications in June 2013, already warned that "*From an enforcement perspective, DoNotTrack could remain a sugar pill instead of being a proper cure and would such be useless*." The Working Group recommended "***the default setting should be such that the user is not tracked***" and that "*In the*

[26] http://www.centerfordigitaldemocracy.org/protect-consumer-privacy
[27] http://www.pubmatic.com
[28] http://www.neustar.biz
[29] http://www.mediamath.com
[30] http://www.exelate.com
[31] http://www.rubiconproject.com/
[32] http://www.appnexus.com/
[33] http://rocketfuel.com/
[34] https://epic.org/privacy/consumer/online_tracking_and_behavioral.html
[35] http://www.ftc.gov/sites/default/files/documents/public_comments/proposed-consent-agreement-matter-google-inc.google-buzz-file-no.1023136-00032%C2%A0/00032-58542.pdf
[36] http://www.centerfordigitaldemocracy.org/cdd-calls-ftc-protect-privacy-todays-hyper-local-geo-targeting-cross-platform-big-data-erawarns-disc
[37] http://www.centerfordigitaldemocracy.org/doc/cdd-testimony-20090618
[38] http://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy/do-not-track
[39] http://www.aboutads.info/
[40] http://www.aboutads.info/how-interest-based-ads-work#what-can-I-do
[41] http://www.aboutads.info/choices/#completed

absence of fully informed user choice, a Web tracking organization must assume that a user is not aware of Web tracking and therefore assume the default position as if they had received a DoNotTrack:1 signal, which indicates a wish from the user not wanting to be tracked". [42]

In July 2013, the World Wide Web Consortium rejected the "DoNotTrack" standard proposed by the online advertising industry (DAA) because the proposal was "*less protective of privacy and user choice than their earlier initiatives*". [43]

The techniques are increasingly sophisticated, **and user tracking is evolving to cross-platform tracking**, so companies can re-target users who switch from desktop to mobile devices. "*Between November 2012 and March 2013, more than 55 million people who visited Expedia's website via their desktop computers and later switched to using mobile devices saw advertisements on their mobile browsers encouraging them to download or use the Expedia mobile app. No other online travel agency has ever re-targeted and tracked ads across device platforms at the same users on a similar scale*"[44]. Expedia[45] uses Drawbridge[46] technology for tracking a user's Web browsing activity and serve relevant advertisements based on cookies or "pixels", which are bits of code usually used to read and place cookies.

"*Very sensitive information is often collected, including health and financial data. One company, Healthline, lets advertisers track people with bipolar disorder, overactive bladder, or anxiety - producing ads related to those conditions targeted at specific people.* ***Advertisers collect, use, and sell social security numbers, financial account numbers, and information about sexual behavior and sexual orientation with no controls or limits****."*[47]

On the 27th of January 2014, the New York Times online published that even the US government was worried about online advertising privacy intrusive practices. "*President Obama announced new restrictions this month to better protect the privacy of ordinary Americans and foreigners from government surveillance, including limits on how the N.S.A. can view the metadata of Americans' phone calls — the routing information, time stamps and other data associated with calls. But he did not address the information that the intelligence agencies get from leaky apps and other smartphone functions.* ***Mr. President Obama expressed concern about advertising companies that collect information on people*** *to send tailored ads to their mobile phones,* ***he offered no hint that American spies have routinely seized that data***"[48]. The article also highlights the fact that **there is no evidence of cooperation of intelligence agencies with advertisement companies in the documents leaked by Edward Snowden**: "*Nevertheless,* ***nothing in the Snowden revealed secret reports indicates that the companies cooperated with the spy agencies to share the information****; the topic is not addressed.*"[48]

The online behaviour tracking techniques and products are not only used by online advertising companies but by other stakeholders too. For instance, the Irregular Times group is developing a *2016 Presidential Campaign Surveillance Report*[49] that analyses the suspected use of such tools by the main identity websites of US politicians taking part in the 2016 presidential campaign, without prior notification to the visitors.

[42] http://www.datenschutz-berlin.de/attachments/949/675.46.13.pdf?1370523228
[43] http://www.w3.org/2011/tracking-protection/2013-july-decision/
[44] http://www.tnooz.com/article/how-expedia-ads-now-trail-you-from-desktop-to-mobile/#
[45] http://expediainc.com/
[46] http://www.drawbrid.ge/
[47] http://online.wsj.com/news/articles/SB10001424052748703940904575395073512989404
[48] http://www.nytimes.com/2014/01/28/world/spy-agencies-scour-phone-apps-for-personal-data.html?_r=0
[49] http://irregulartimes.com/2016surveillance.html

1.3 Question 3

The use of cookies and trackers shows its limits due to generalisation of the use of sophisticated cryptography and the explosion of a number of connected devices (Internet of everything).
How are data trade organisations such as the ones cited under (2) above and mass surveillance organisations preparing for "Privacy and Tracking in a Post-Cookie World"?
Are there likely impacts in terms of additional hooks and backdoors implemented by software editors for commercial purposes in operating systems and applications?
How could end-users be able to protect themselves from such more aggressive personal data collection techniques in the future?
Is this feasible?
Is this desirable?

ENISA's work on cookies[50] provides a comprehensive explanation on what cookies are, how they work and the privacy and security concerns they raise. In summary, **cookies are text files that are placed in the web browser by the websites visited by the user. Cookies allow for smarter and faster navigation, and are commonly used for personalizing website content as well as ads and features by associated third parties** with which the website collaborates to offer more engaging content. **Cookies allow for state management over the HTTP** (Hypertext Transfer Protocol) protocol which is stateless by nature, i.e. the response-request pairs in the communication are treated as independent transactions not related to other pairs. Therefore, any request from the client would not be correlated to a previous request which is necessary for a fluent navigation experience where the user is not continuously asked to enter the same data (e.g. for authentication purposes).

In order to overcome this situation, the cookies were created as part of Internet standards to keep state information. The **cookies are generated and modified by website servers, stored in client's browsers and transmitted in every interaction between the server and the browser**.

In terms of life-span, **there are two main types of cookies: session cookies that are temporarily stored in memory and deleted once the connection session times out** or when the browser is closed; **and persistent cookies that span over sessions and remain in the web browser, even when the browser is closed**, till their expiration date. While session cookies' purpose is to keep state information within sessions, persistent cookies are used for relating subsequent sessions or visits to a website.

The information that can be stored in and extracted from the cookies is diverse and ranges from user registration and log in information (e.g. credentials and other identifiers), user preferences and settings, session data, data cached by the site, records of user browsing activity, and other information that can be retrieved in future sessions, including personally identifiable information provided to the website by the user.

With such a wide variety of possibilities, it is not surprising that advertisement and surveillance organisations are exploiting the cookies to collect and compile tracking information about online users and to learn about their browsing histories and online behaviour. With the advent of new techniques and new market needs, a plethora of new cookie types (e.g. supercookies, flash cookies, etc.) have been created by marketing and commercial industry to optimize the aggregation of data and inference

[50] http://www.enisa.europa.eu/activities/identity-and-trust/privacy-and-trust/library/pp/cookies

of user tracking information. All these types are very well explained in the literature, for example ENISA work on cookies and Wikipedia[51] itself do explain the classes in easily understandable terms.

For instance, in November 2014 "The Washington Post" published that Verizon and AT&T have been tracking the Internet activity of more than 100 million customers with supercookies[52].

Websites today are increasingly complex and orchestrate a large amount of external content and services from third party systems and vendors from businesses such as advertising, social networking, analytics, etc. **The proliferation of third party cookies and the ever-growing myriad of IT platforms or devices from which users connect to Internet have made the cookies an insufficient solution.** Since cookies are specific to a website or domain, each third party issues at least one third-party cookie per domain, which in turn slows down website's load times and causes concerns of "data leakage".

The **lack of transparency and control over their data are the main reason why users have started to block or delete cookies** and to adopt best practices for increased privacy, such as the FTCs recommendation for a "Do Not Track" mechanism[53].

Following this trend of respecting consumers concerns, **third-party cookies are not always accepted and** even **various browsers**, for example mobile Safari, which is used on most iPhones and iPads, **avoid their use in the default settings**. Others such as Firefox are considering blocking all third party cookies. **Cookies have** also **limited utility for collecting data across applications or devices** as they are specific to a unique login session, a unique device, and a unique browser or application. This limits the capability of cookies to master user location data and keep track of the user identity across platforms.

For these reasons, **the use of traditional cookies is now being superseded by other state management solutions that overcome cookies limitations.** The Future of the Cookie Working Group[54] of Interactive Advertising Bureau (IAB) is working in pushing cookie-free solutions that can address simultaneously all requirements of the three involved players': consumers, publishers (creators, facilitators and/or owners of website contents) and third parties.

In their whitepaper "Privacy and Tracking in a Post-Cookie World"[55], IAB explores a set of four different technologies that could replace the cookies, and which address the three stakeholder types' needs with different degrees of success. The one page infographic summary of the solution classes[56] (see Figure 2) is explicative enough, and compares the traditional cookies (Server-Issued State solution) with the other four alternative technologies. The work defines the solutions as follows:

- **Device-Inferred State** – State managed through the use of IDs inferred using statistical algorithms applied to information passed by the device, browser, app or operating system.
- **Client-Generated State** – State and preferences managed from within the client (such as the browser, app, or operating system) and passed to third parties within the ecosystem. Examples in the market today include the Advertising ID on iOS and Android.
- **Network-Inserted State** – State and preferences managed via IDs set by third party intermediary servers that sit between the end-consumer's device and the publishers' servers. Examples include

[51] http://en.wikipedia.org/wiki/HTTP_cookie#Supercookie

[53] Federal Trade Commission. (2010, December 01). FTC Staff Issues Privacy Report, Offers Framework for Consumers, Businesses, and Policymakers. Retrieved from ftc.gov: http://www.ftc.gov/opa/2010/12/privacyreport.shtm

[54] http://www.iab.net/member_center/committees/working_groups/Future_of_the_Cookie_Working_Group

[55] http://www.iab.net/media/file/IABPostCookieWhitepaper.pdf

[56] http://www.iab.net/futureofcookie

content distribution networks, Wi-Fi or wireless proxy servers and ISPs. This is a concept not broadly offered in the market today.

- **Server-Issued State** – State and preferences managed via HTTP cookies set between each server domain and browser client, often via Web beacons or pixels. This is the incumbent approach which is ubiquitous in the market today.
- **Cloud-Synchronized State** – State and preferences managed via IDs set by a centralized service that all parties agree to work with. This is a concept not broadly offered in the market today.

Other recent work by Marc Groman[57] from Network Advertising Initiative (NAI) proposes three possible categories of alternatives to HTTP cookies:

- **Statistical IDs** – the use of statistical algorithms that use information passed by the device, browser or operating system to infer a user ID that can then be used by publishers or third parties.
- **Client IDs** – deterministic identifiers that sit on the app or browser. Some are set by the server such as HTML5 local storage and so-called "flash cookies." Others are created client side by the OS such as the platform IDs on Androids and iOS.
- **Centralized Cloud-Synchronized IDs** – IDs set and managed through a centralized cloud service that all parties in the ecosystem agree to work with instead of a third party intermediary such as an ISP.

Although the first two options do already exist in the market, the Cloud version is the most innovative and still being explored by emerging works such as DigiTrust[58]. The last two of these technologies can be directly associated to the respective Client-Generated State and Cloud-Synchronized State solution classes proposed by IAB.

[57] https://privacyassociation.org/media/presentations/14Academy/IAPP-PPT-Revised_Groman.pdf at International Association of Privacy Professionals (IAPP) Privacy Academy and CSA Congress 2014.
[58] http://www.digitru.st/

Each solution class is evaluated against each guiding principle as follows:

- The state management solution class includes core capabilities that support the principle.
- The state management solution class may support the principle in certain implementations.
- The state management solution class has significant challenges to overcome in order to support the principle.

	Guiding Principle	Server Solution	Device Solution	Client Solution	Network Solution	Cloud Solution
Consumer	Single privacy dashboard	🔴	🟡	🟡	🟡	🟢
	Universal privacy view	🔴	🟡	🟡	🟡	🟢
	Comprehensive privacy controls	🟢	🟡	🟢	🟢	🟢
	Persistent, universal consumer preferences	🔴	🟡	🟡	🟡	🟢
	Possibility of detecting non-compliant actors	🔴	🔴	🔴	🔴	🟡
	Free online service	🟢	🟢	🟢	🟢	🟢
Publisher	Single privacy dashboard	🔴	🟡	🟡	🟡	🟢
	Comprehensive privacy controls	🔴	🟡	🟡	🟡	🟢
	Significantly fewer 3rd party pixels	🔴	🟢	🟢	🟢	🟢
	Improved user tracking/targeting	🔴	🟡	🟢	🟢	🟢
	Reduced cost for privacy compliance	🔴	🟡	🟢	🟢	🟢
	Certified participant visibility	🔴	🟡	🟢	🟢	🟢
	Doesn't tilt towards a specific vendor	🟢	🟡	🟡	🟡	🟡
	Minimal deployment overhead	🟢	🟡	🟡	🟡	🟡
Industry	Decreased segment ramp-up time	🔴	🟢	🟢	🟢	🟢
	Decreased "cookie churn"	🔴	🟡	🟢	🟢	🟢
	Lower operating cost	🔴	🟡	🟡	🟡	🟡
	Better cross-device tracking	🔴	🟡	🟡	🟡	🟢
	Better consumer transparency/control	🔴	🔴	🟢	🟢	🟢
	Higher integrity frequency capping	🔴	🟢	🟢	🟢	🟢
	Less redundant data collection/transfer	🔴	🟡	🟢	🟢	🟢
	Reduced regulatory threats	🔴	🟡	🟢	🟢	🟢
	Clarifies value to consumer	🟡	🔴	🟡	🟡	🟡
	Doesn't tilt towards a specific vendor	🟢	🟡	🟡	🟡	🟡
	Minimal deployment overhead	🟢	🟡	🔴	🔴	🔴

Figure 2: IAB proposal of solution classes to replace traditional cookies

Whatever the **future alternative to cookies**, in order to be consistent with existing regulatory and self-regulatory standards, they **need to ensure transparency to Internet users on which data is stored or transmitted to the website, as well as which types of statistical analysis is performed with such data**, for what purposes and with whom it is shared. In addition, new solutions need to let users choose to opt-out for websites not tracking their data, similar to opt-out mechanisms that current cookie technologies provide. **In most of the cases, current solutions still lack browser plugins or other mechanisms for such transparency or control**.

As highlighted by IAB, one of the most important factors for the cookie alternatives to be widely adopted by the Internet industry and its trustworthiness ensured, is the need of being conceived and deployed as open solutions. The **technologies shall neither be proprietary nor licensed by** one or a few **commercial organisations**. They should but be as openly accessible as possible, if not directly embedded in Internet standards that are widely supported by public authorities all over the globe.

In such a future landscape, the mass surveillance organisations would have great difficulties in tracking users without consent unless they illegally collaborate with commercial organisations to

implement hooks and backdoors in operating systems and applications. It is not easy to anticipate to what degree such illegal activities could become a reality.

1.4 Question 4

> ***Are mass surveillance organisations limiting themselves to exploiting information provided by commercial cookies and trackers such as "Google Analytics" for instance?***
> ***Is there some technical evidence that mass surveillance organisations cooperate with commercial websites or even possibly hack them, to deploy their own set of trackers and cookies over commercial website infrastructures? What type of personal data can they succeed to gather that way?***
> ***What is the limit of what can be spied this way?***

The summary of the more complete catalogues on currently used **mass surveillance resources** (tools, programs, hardware, etc.) **is provided in the answer to Question 19**. These catalogues show the huge potential for personal data collection, storage and analysis that mass surveillance organisations do have.

Although it is true that some of those tools were developed by governmental security agencies and are not available for commercial purchase, many others are publicly promoted and sold in Internet. **These tools** serve many different purposes and operate in many different ways, but they certainly **go far beyond the mere tracking and use of commercial cookies**.

Commercial cookies are a mechanism for websites to remember stateful information that the HTTP protocol by itself does not provide. The tracking cookies and especially third-party tracking cookies are commonly used as ways to compile long-term records of individuals' browsing histories, as **cookies can expose a number of aspects from the website visitors, such as the user's log-in information, browsing events activity (including clicking particular buttons, ads, etc.), browser attributes and user preferences, passwords (usually stored encrypted), etc**. The long term recording and later analysis of such information serves for **profiling of user on-line behaviour**, which **is commonly exploited by commercial websites to customize the ads, application menus, and other information** for their clients.

Google Analytics[59] and trackers of the like, such as Piwik[60], KISSMetrics[61] and Clicky[62], provide detailed information on website traffic statistics in real time and other advanced functionality like event and funnel tracking or segmenting for filtering data.

These data can reveal users' likes and dislikes, browsing habits and trends, etc. and their tracking has raised concerns on privacy protection. Both, European and U.S. regulatory frameworks intensively started to study the problem since 2011. But collections of much more sophisticated tools exist, with which mass surveillance organisations do not even need to cooperate with commercial websites, as they are able to directly hack them. One of the most outstanding examples of such tools that allow direct hacking is the FinFisher spyware, sold by U.K. based Gamma Group, which can secretly take remote control of a computer, copy and exfiltrate files, intercept Skype calls, and log every keystroke of a computer user[63].

[59] http://www.google.com/analytics/
[60] http://piwik.org/
[61] https://www.kissmetrics.com/
[62] http://clicky.com/
[63] https://citizenlab.org/2012/07/from-bahrain-with-love-finfishers-spy-kit-exposed/

With this kind of tools, the internet user is so vulnerable that any single word he writes or says over Skype can be known and misused by the surveyor. The risk is even greater as **publicly sold spyware is now spreading to the mobile world**. According to a study from the University of Toronto Munk School of Global Affairs' Citizen Lab[64] a number of variants of the FinFisher toolkit exist that can infect BlackBerrys, iPhones, and other mobile devices.

The Finfisher infection starts via a malicious email with an attachment containing malware. Upon execution of the attachment, the malware installs a multi-featured trojan on the victim's computer or handheld device. In the mobile version, when the user gets the malicious email or text message and clicks the included link, the page that loads drops malicious code that pops up a message to update an apparently harmless application (e.g. Apple's iTunes media player) which is actually a fake application that disguises the trojan. If the user executes the update, the spyware application is installed and the remote system can then secretly turn on and record from the device's microphone, track locations, and monitor e-mails, text messages and voice calls.

Another example of spyware that is particularly targeting mobile phone devices is StealthGenie, an app that was sold through Google's AppStore and that is designed to secretly tap all communications (voice, text, chat) on the devices it is installed. StealthGenie was only recently taken off the market, due to an indictment of the US Justice Department[65].

As can be seen in the documentation associated to the tools in the catalogues, most of these tools do not only exploit content but also metadata.

Professor Edward Felten[66] warns that *the telephony metadata can be extremely revealing both, at the level of individual calls and especially in the aggregate.* ***Analysis of metadata often reveals information that could traditionally only be obtained by examining the content of the communication***[67].

Aggregated metadata –i.e. metadata about a single person over time, about groups of people, or bounded with other datasets - "***generates a precise, comprehensive record*" of people's habits, which in turn "*reflects a wealth of detail about [their] familial, political, professional, religious, and sexual associations***[68],[69]"

This revelatory nature of the metadata was also evidenced in a study by Jonathan Mayer and Patrick Mutchler on mobile telephony metadata. It proved that **phone call patterns revealed highly sensitive information about the people being tracked** such as medical conditions and ownership of a specific brand of firearm[70]. This information was inferred with the sole use of Android smartphone logs and Facebook social network information for the analysis.

The study reveals the need for protection of cookies and other logged information in the devices used to connect to Internet. **Unprotected cookies are open doors to a multitude of attacks** such as session fixation (*the attacker fixes the user's session ID before the user even logs into the target server*[71], which allows

[64] Morgan Marquis-Boire and Bill Marczak, "The SmartPhone Who Loved Me: FinFisher Goes Mobile?" https://citizenlab.org/2012/08/the-smartphone-who-loved-me-finfisher-goes-mobile/

[65] http://www.washingtonpost.com/business/technology/make-of-app-used-for-spying-indicted-in-virginia/2014/09/29/816b45b8-4805-11e4-a046-120a8a855cca_story.html

[66] Professor of Computer Science and Public Affairs , as well as Director of the Center for Information Technology Policy , at Princeton University

[67] https://www.aclu.org/files/pdfs/natsec/clapper/2013.08.26%20ACLU%20PI%20Brief%20-%20Declaration%20-%20Felten.pdf

[68] https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2013cv0851-48

[69] https://www.eff.org/files/2014/03/13/clapper_amicus_computer_scientists.pdf

[70] Jonathan Mayer & Patrick Mutchler, MetaPhone: The Sensitivity of Telephone Metadata (Mar. 12, 2013), http://webpolicy.org/2014/03/12/metaphone-the-sensitivity-of-telephone-metadata/

[71] http://www.acros.si/papers/session_fixation.pdf

the attacker impersonate the user), cookie-based SQL injection[72], cookie-stealing attacks to gain access to private information such as web accounts, etc.

In December 2013, The Washington Post reported[73] that **NSA was suspected of using Google cookies to identify targets for hacking their computers**. The way these tracking cookies were gained access to is not addressed in the article. *It is theorized that one way the NSA could get access to the tracking cookies is to simply ask the companies for them under the authority granted to the agency by the Foreign Intelligence Surveillance Act (FISA)[74].*

1.5 Question 5

> ***What type of report or information can be practically produced and consolidated over a long period of time by analysing meta-data on a mass scale at a country level for instance?***
> ***Are Big Data and Big Analytics type technologies used for mass surveillance purposes, in order to consolidate useful information on a large scale?***
> ***How potentially privacy intrusive can be the information collected and synthetized that way?***

Todays increased data collection and novel approaches for data representations and mathematical modelling coincide with the development of powerful database technologies that provide an easy access to the massive amounts of collected data[75]. These include technologies to deal with non-structured data as well as structured data.

"Big Data" is a term encompassing the use of techniques to capture, process, analyse and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies. By extension, the platform, tools and software used for this purpose are collectively called "Big Data technologies"[76]. This type of technologies deals with huge amounts of distributed and heterogeneous data such as structured data (e.g. metadata) and non-structured data (e.g. voice). Furthermore, data collection and analysis is being conducted at a velocity that is increasingly approaching real time[77].

The structured nature of the metadata is ideally suited for analysis using data mining techniques such as pattern recognition, machine learning, and information or data fusion. These knowledge-discovery-in-databases (KDD) techniques are rapidly evolving and providing new innovative tools such as support for vector machines, genetic algorithms, classification and regression trees, Bayesian networks, and hidden Markov models, to make better use of this explosion of information[75].

As pointed out by Professor Edward Felten[78] it is difficult to summarize all the sensitive information that can be revealed by processing only metadata (not content). But suffice it to say that **it can reveal an extraordinary amount of information about people's habits and associations.**

As it is explained by Felten: "*Calling patterns* [deduced from metadata] *can reveal when we are awake and asleep; our religion, if a person regularly makes no calls on the Sabbath, or makes a large number of calls on*

[72] http://resources.infosecinstitute.com/cookie-based-sql-injection/
[73] http://www.businessinsider.com/how-the-nsa-uses-cookies-to-hack-computers-2013-12
[74] http://www.computerworld.com/article/2486724/mobile-security/nsa-taps-tracking-cookies-used-by-google--others--to-monitor-surveillance-targets.html
[75] http://epic.org/misc/nrc_rept_100708.pdf
[76] Big Data – A new world of opportunities, NESSI White Paper, December 2012
[77] Big Data: Seizing opportunities, preserving values. US Government report. May 2014.
[78] Professor of Computer Science and Public Affairs , as well as Director of the Center for Information Technology Policy , at Princeton University

Christmas Day; our work habits and our social aptitude; the number of friends we have; and even our civil and political affiliations."

When metadata is aggregated – data over time, or linked with other datasets - it can expose even richer personal information and associational details. Analysis of this kind of metadata can reveal the network of individuals with whom they communicate, commonly called a "social graph". "*Data-mining systems for national security use are designed to link any common identifying numbers of any kind and look for correlations, geographical intersections of location data, and patterns in online social relationships.* ***Unless special precautions are taken, few personal secrets of everyday life would withstand close analysis of metadata".***[79]

Big data technologies are critical for mass surveillance purposes since they provide tools to easily process and analyse vast amount of heterogeneous data. For instance, the NSA created a computer software database project called "Accumulo" in 2008, which can be classified as Big Data technology and submitted it to the Apache Foundation in 2011 as an incubator open-source project[80].

Snowden's revelations have also shown the use of Big Data technologies. "Boundless Informant" is a Big Data analysis and data visualization tool developed and used by the NSA. The objective of this tool is to count and categorize the metadata of both, internet and telephony communications[81].

1.6 Question 6

What is the level of technical complicity required from the different network infrastructure operators and commercial websites?
Can surveillance be done independently of their knowledge by "listening network infrastructure secretly" and by "hacking commercial websites" to deploy mass surveillance cookies, trackers and other malware?
How can citizens detect that their meta-data is being analysed for mass surveillance purposes and how can citizens protect themselves accordingly?

Secret US Government documents, leaked by Edward Snowden and published by the media in 2013, **confirm the US intelligence services have obtained copies of the internet traffic that is transmitted through the US major domestic fibre-optic cable networks**[82,83]. NSA has its own cable-intercept programs for internet traffic surveillance purposes, which operate mainly under four codenames - BLARNEY, FAIRVIEW, OAKSTAR and STORMBREW- and are collectively known as Upstream collection[84]. In the case of the US, the data collection is conducted under different Law Authorities: Transit Authorities, Foreign Intelligence Surveillance Act (FISA) and FISA Amendment Act of 2008 (FAA)[85].

[79] http://blog.privacystrategy.eu/public/published/Submission_ISC_7.2.2014_-_Caspar_Bowden.pdf
[80] http://www.informationweek.com/applications/nsa-submits-open-source-secure-database-to-apache/d/d-id/1099972
[81] http://www.theguardian.com/world/2013/jun/08/nsa-boundless-informant-global-datamining
[82] https://www.eff.org/nsa-spying
[83] http://www.washingtonpost.com/world/national-security/nsa-tracking-cellphone-locations-worldwide-snowden-documents-show/2013/12/04/5492873a-5cf2-11e3-bc56-c6ca94801fac_story.html
[84] http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#section/3
[85] http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#doc/2

In June 2013, the media also reported that **the British intelligence agency GCHQ intercepts, collects and stores data from the fibre-optic cable network which carries international phone and internet traffic data**[86].

Internet monitoring is the act of intercepting data packets over Internet Protocol. The infrastructure that supports the Internet includes physical infrastructure and electronic systems (such as communication switches, routers, servers, etc.) to connect the world. **Internet monitoring can take place across any point of this infrastructure**, depending on what information shall be collected[87]. It is interesting to know that the US is connected to 63 countries by fibre-optic cables; France to 60; Portugal to 59; UK, to 57; Italy to 47; Spain to 41; Germany to 40; Greece to 37[83]. Therefore the volume of internet data that is transmitted through these fibre-optic cables is huge and is susceptible to be intercepted.

There are other cases in which telecom operators and internet organizations under repressive regimes have bought, installed and maintained software and equipment for phone and internet traffic interception[88]. In case of Tunisia, the German company Trovicor GmbH provided voice and data interception on cell phones[89].

The ACLU published a schematic diagram representation of the NSA surveillance over the major communication switches, routing stations, and access points of telecommunication systems[90]. According to ACLU, this type of data collection includes both, access to gateways through which phone calls are routed, as well as access to other telecommunication equipment (such as routers, switches, etc.) through which internet traffic data passes.

Vodafone revealed the existence of secret wires that allow government agencies to intercept all conversations on its networks under the law in place. This practice is widely used in some of the 29 countries in which Vodafone operates in Europe and beyond[91].

However, according to the German magazine "Der Spiegel", the NSA has a special unit comprised of agents specialized in backdoors and malware software and they are **able to build tools for penetrating networking equipment, monitoring mobile phones and computers and diverting or even modifying data**[92]. Data interception (either by the NSA or any other surveillance organization) using **this technique implies that the organization which has been attacked is not aware of it and the data collection can be done without any collaboration from their side**.

But data collection is also carried out by requesting telecommunication operators or internet agencies to send phone and internet related data to the intelligence agencies under a law order, as happened in the Verizon case[93,94]. As indicated in the white paper released by the US Government, "*under the telephony metadata collection program, **telecommunications service providers, as required by court orders issued by the FISC, produce to the Government certain information about telephone calls**, principally those made within the United States and between the United States and foreign countries*"[95,96].

[86] http://www.theguardian.com/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa
[87] https://www.privacyinternational.org/sii/technologies/internet-monitoring
[88] http://www.wired.com/2011/01/as-egypt-tightens-its-internet-grip-tunisia-seeks-to-open-up/
[89] http://www.bloomberg.com/news/2011-12-12/tunisia-after-revolt-can-alter-e-mails-with-big-brother-software.html
[90] https://www.aclu.org/files/pdfs/eavesdropping101.pdf
[91] http://www.theguardian.com/business/2014/jun/06/vodafone-reveals-secret-wires-allowing-state-surveillance
[92] http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994.html
[93] http://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order
[94] http://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded#doc/1
[95] http://www.theguardian.com/world/2013/aug/09/obama-legal-background-surveillance-nsa

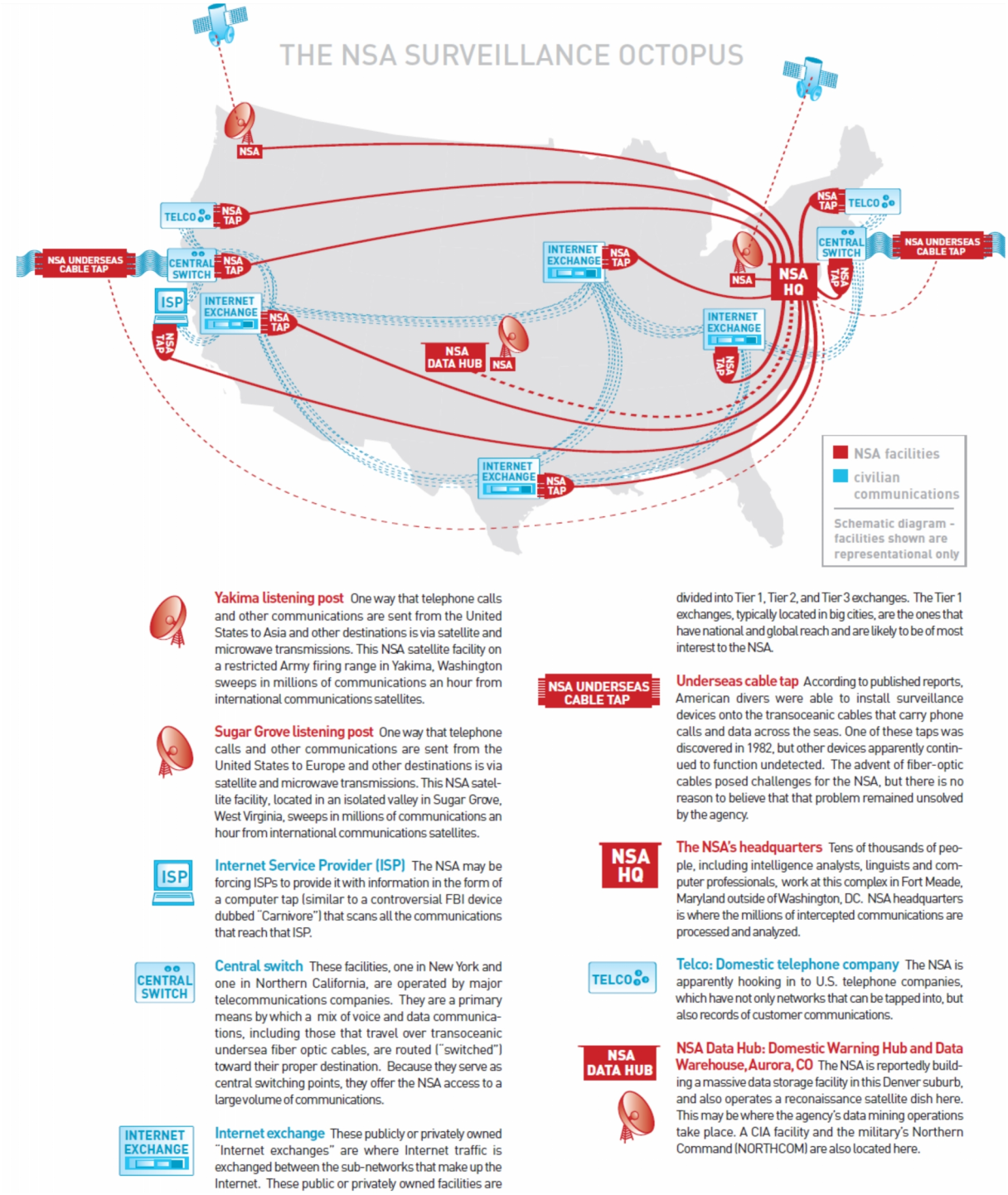

THE NSA SURVEILLANCE OCTOPUS
NSA
TELCO
NSA TAP
NSA UNDERSEAS CABLE TAP
CENTRAL SWITCH
ISP
INTERNET EXCHANGE
NSA DATA HUB
NSA HQ
NSA facilities
civilian communications
Schematic diagram - facilities shown are representational only
Yakima listening post One way that telephone calls and other communications are sent from the United States to Asia and other destinations is via satellite and microwave transmissions. This NSA satellite facility on a restricted Army firing range in Yakima, Washington sweeps in millions of communications an hour from international communications satellites.
Sugar Grove listening post One way that telephone calls and other communications are sent from the United States to Europe and other destinations is via satellite and microwave transmissions. This NSA satellite facility, located in an isolated valley in Sugar Grove, West Virginia, sweeps in millions of communications an hour from international communications satellites.
Internet Service Provider (ISP) The NSA may be forcing ISPs to provide it with information in the form of a computer tap (similar to a controversial FBI device dubbed "Carnivore") that scans all the communications that reach that ISP.
Central switch These facilities, one in New York and one in Northern California, are operated by major telecommunications companies. They are a primary means by which a mix of voice and data communications, including those that travel over transoceanic undersea fiber optic cables, are routed ("switched") toward their proper destination. Because they serve as central switching points, they offer the NSA access to a large volume of communications.
Internet exchange These publicly or privately owned "Internet exchanges" are where Internet traffic is exchanged between the sub-networks that make up the Internet. These public or privately owned facilities are divided into Tier 1, Tier 2, and Tier 3 exchanges. The Tier 1 exchanges, typically located in big cities, are the ones that have national and global reach and are likely to be of most interest to the NSA.
Underseas cable tap According to published reports, American divers were able to install surveillance devices onto the transoceanic cables that carry phone calls and data across the seas. One of these taps was discovered in 1982, but other devices apparently continued to function undetected. The advent of fiber-optic cables posed challenges for the NSA, but there is no reason to believe that that problem remained unsolved by the agency.
The NSA's headquarters Tens of thousands of people, including intelligence analysts, linguists and computer professionals, work at this complex in Fort Meade, Maryland outside of Washington, DC. NSA headquarters is where the millions of intercepted communications are processed and analyzed.
Telco: Domestic telephone company The NSA is apparently hooking in to U.S. telephone companies, which have not only networks that can be tapped into, but also records of customer communications.
NSA Data Hub: Domestic Warning Hub and Data Warehouse, Aurora, CO The NSA is reportedly building a massive data storage facility in this Denver suburb, and also operates a reconaissance satellite dish here. This may be where the agency's data mining operations take place. A CIA facility and the military's Northern Command (NORTHCOM) are also located here.

Figure 3: Schematic diagram of NSA's interception nodes

But **mass surveillance organizations can intercept data without the collaboration of the network operators** or commercial website's operators. However the techniques utilized are more targeted than the mechanisms explained before. Some of the techniques that don't need collaboration are described below.

[96] https://www.documentcloud.org/documents/750211-administration-white-paper-section-215.html

"Stingray device", new location tracking device. The so-called "stingray" is a device that can be used to track a suspect's location and engage in other types of surveillance[97]. The more generic term for the device is "IMSI catcher"[98]. **IMSI catchers imitate a wireless carrier's network equipment and** in doing so, they send and receive signals to and from all mobile devices in the local area on the same network[96]. This surveillance tool **can collect information about the target devices as well as locations of third parties**.

HACIENDA port scanning program. In august 2014, "Heise" revealed, co-written by Jacob Appelbaum and Laura Poitras, the existence of **a mass surveillance technology that scans open ports on all servers connected to the Internet searching for vulnerabilities** to be exploited[99]. According to this new article, **this program was built by the GCHQ**.

Malware/Spyware. Malware is any malicious software, script, or code run on a device that alters its state or function without the owner's informed consent[100]. Malware comes in many different forms of which viruses, worms and trojans are used most often. Spyware is malware that is specifically designed to collect, monitor, and log the actions of a system user. Thus, this type of **software** is **capable of penetrating networking equipment, monitoring mobile phones and computers and diverting or even modifying data**[101].

"FinSpy" malware, part of the commercial intrusion kit "Finfisher" distributed by the Gamma Group, collects and encrypts a wide range of data from the infected device. The "FinSpy Mobile" provides recording of common communications such as voice calls, emails and SMS/MMS, file download from the infected device and location tracking[102]. Furthermore, this malware **employs a myriad of techniques designed to evade detection and resists any kind of analysis**[103].

For example, a **surveillance organization can produce and publish fake copies of most popular websites on the Internet** such as LinkedIn and once the user has opened the website a malware is installed on the computer and the mass surveillance organization can gain access to the user's networks. According to "Der Spiegel", **GCHQ created a fake LinkedIn website in order to collect information related to mobile communication companies and billing companies by targeting their employees**[104]

*In some cases the **NSA has masqueraded as a fake Facebook server, using the social media site as a launching pad to infect a target's computer** and exfiltrate files from a hard drive. In others, it has sent out spam emails laced with the **malware,** which **can be tailored to covertly record audio from a computer's microphone and take snapshots with its webcam**. The hacking systems have also enabled the NSA to launch cyberattacks by corrupting and disrupting file downloads or denying access to websites*"[105].

Hacking internet forums can be another example, as explained in an article that was published on the Dutch news-site NCR: ***The Dutch intelligence service** - AIVD - **hacks internet web forums to collect the***

[97] https://www.aclu.org/blog/national-security-technology-and-liberty/court-uncovering-stingrays-troubling-new-location

[98] http://www.emsec.rub.de/media/crypto/attachments/files/2011/04/imsi_catcher.pdf

[99] http://www.heise.de/ct/artikel/NSA-GCHQ-The-HACIENDA-Program-for-Internet-Colonization-2292681.html?hg=1&hgi=3&hgf=false

[100] Verizon, VERIS (Vocabulary for Event Recording and Incident Sharing).

[101] http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994.html

[102] Morgan Marquis-Boire, Bill Marczak, Claudio Guarnieri and John Scott-Railton. For Their Eyes Only: The Commercialization of Digital Spying. Citizen Lab and Canada Centre for Global Security studies, University of Toronto. May 2013.

[103] https://citizenlab.org/2012/07/from-bahrain-with-love-finfishers-spy-kit-exposed/

[104] http://www.spiegel.de/international/world/ghcq-targets-engineers-with-fake-linkedin-pages-a-932821.html

[105] https://firstlook.org/theintercept/2014/03/12/nsa-plans-infect-millions-computers-malware/

data of all users. *The majority of these people are unknown to the intelligence services and are not specified as targets when the hacking and data-collection process starts"*[106].

For an end user it is practically impossible to detect whether meta-data generated while navigating through the web, sending mails, or establishing other communications through internet **is being analysed or used by third parties**, with the exception of the use of cookies by 3rd party advertisers, which can be observed through the placement of "customised" ads on web pages. Both, malware technology and IP monitoring technology vendors claim their tools are invisible to the target: "*Remote Control System" is totally invisible to the target. It bypasses protection systems such as antivirus, antispyware and personal firewalls*[107]"; "*… POSEIDON is totally invisible in a communication network and cannot be identified*[108]. Or as Bruce Schneier, a renowned security and cryptography expert, puts it in an article published by the "Guardian": "[NSA's TAO has a] *variety of tricks to get* [exploits] *on to your computer. Your anti-virus software won't detect them, and you'd have trouble finding them even if you knew where to look*"[109]

The perfect means for a citizen to prevent analysis of personal meta-data is simply not producing it, which essentially means not using electronic communication mechanisms. This is of course not a feasible alternative and the situation can therefore only be addressed by practices and solutions that would help to reduce, or hide, the generated metadata.

There are a number of available techniques that citizens can use to protect their privacy online. The techniques to use depend on the type of communication, the device and platform used for communication and the life-cycle of data that shall be protected.

One of the first and most obvious steps for achieving this is to **restrict the use of cookies in the browser settings**. Most browsers include options for "stealth" navigation, which impede the storage of navigation data (images, text, cookies, history, etc.). This does, however, influence the user experience in the sense that, for instance, preferences for a particular page or services which are normally stored in cookies cannot be maintained. It must also be said that **this option offers virtually no privacy protection beyond the local level**, because visited websites can for example be identified by matching a user's IP address on the server end.

Even if full protection is not possible and agencies such as NSA are suspect of having tried to break some encryption algorithms[110], cryptography experts[111] do still recommend using encrypted communications for the protection of the meta-data. Even the whistle-blower Edward Snowden stated that: "*properly implemented* ***strong crypto systems are one of the few things you can rely on***"[112]. Similarly, ProPublica[113] journalist Julia Angwin declares: "*ProPublica has written about the NSA's attempts to break encryption, but we don't know for sure how successful the spy agency has been, and security experts still recommend using these techniques*"[114]. Her article recommends using encryption for user data both, at transfer and at rest.

The following list summarizes the most recommended tools and means for keeping user identity, user content or metadata safe from intrusive eyes. It follows the classification of data in transfer and data at rest.

---

[106] http://www.nrc.nl/nieuws/2013/11/30/dutch-intelligence-agency-aivd-hacks-internet-fora/
[107] https://www.privacyinternational.org/sii/companies/Hacking_Team
[108] https://www.privacyinternational.org/sii/document/441689-poseidon-ip-information
[109] http://www.theguardian.com/world/2013/sep/05/nsa-how-to-remain-secure-surveillance
[110] http://www.nytimes.com/2013/09/06/us/nsa-foils-much-internet-encryption.html?_r=0
[111] http://www.technologyreview.com/news/519171/nsa-leak-leaves-crypto-math-intact-but-highlights-known-workarounds/
[112] http://www.theguardian.com/world/2013/jun/17/edward-snowden-nsa-files-whistleblower?CMP=twt_gu#block-51bf3588e4b082a2ed2f5fc5
[113] http://www.propublica.org/
[114] http://www.propublica.org/article/privacy-tools-encrypt-what-you-can

Encryption of data at rest (stored data)

Data encryption is a means for protecting data confidentiality and avoiding unauthorised access to the data by parties who do not possess the original encryption keys.

- **Hard drive encryption:** The technique consists in either encrypting entire partitions of a hard drive, or only individual files stored in a partition. Examples of tools that allow to do so are the following:
    - **DiskCryptor**[115]: DiskCryptor is an open-source encryption solution for Microsoft Windows that offers encryption of the entire computer hard drive or individual disk partitions (including the system partition), or external storage devices (USB, DVD disks, etc.). DiskCryptor provides a broad choice in configuration of booting an encrypted operating system (OS). It uses AES-256, Twofish, and Serpent, or combinations thereof to carry out encryption.
    - **TrueCrypt**, is a free and open-source program which was very popular, but has been **discontinued** in May 2014. TrueCrypt supported Microsoft Windows, OS X and Linux, and allowed on-the-fly encryption of individual files, entire hard drives, entire partitions, or storage devices such as a USB flash drives or external hard drives. To this aim three different algorithms were available: AES, Serpent, and Twofish, and five different cascaded combinations of them. There are two forks of the original source code that claim to have taken over the development of the future TrueCrypt: https://truecrypt.ch/ (TCnext) and https://ciphershed.org/.
    - **FileVault**[116]: Is an Apple Macintosh built-in encryption system that is installed on most recent Mac computers. The current FileVault 2, which works with OS X Lion or later, uses full disk, XTS-AES 128 encryption to allow on-the-fly encryption (real time, i.e. as data are stored) of the contents of the entire drive. FileVault 2 requires that OS X Recovery is installed on the startup drive.
    - **BitLocker**[117]: Is a Microsoft built-in encryption system that works in Ultimate and Enterprise editions of Windows 7 and allows encrypting entire drives.
    - **PGP**[118] is an encryption application that was first developed by Phil Zimmerman in 1991 and has evolved under different owner companies. In 2010 PGP has been acquired by Symantec and is now only available through this company. However, an implementation of openPGP is available as freeware from the GnuPG[119] website.

- **Encryption of smartphone's hard drive:** Apple doesn't let users encrypt their smart phone's hard drive or the files on it, though the operating system (< iOS 8) will encrypt passwords and some other files if a passcode is used on the device. Apple will also let users encrypt their phone's backup files on iTunes or iCloud. It is also possible to use Find my iPhone to remotely "wipe," or delete the data on an iPhone or iPad if it is lost or stolen. With the rollout of iOS 8 all data on the phone is encrypted by default. Google's Android operating system lets users encrypt their phone hard drive.
- **Encryption of data in the cloud.** These technologies enable data to be stored encrypted in the cloud, using an encryption key that is owned by the data owner and usually stored on the hard drive of the device accessing the cloud. All files are safely encrypted on the user's device before being transferred to the cloud.

[115] https://diskcryptor.net/wiki/Main_Page
[116] http://support.apple.com/en-us/HT4790
[117] http://windows.microsoft.com/en-us/windows7/products/features/bitlocker
[118] http://www.symantec.com/products-solutions/families/?fid=encryption
[119] https://gnupg.org/

The cloud storage services SpiderOak[120], Wuala[121], BoxCryptor[122] , Cloudfogger[123], Seafile[124] (open source), SparkleShare[125] (open source), and Pydio[126] (open source) are examples of such technologies. The mechanism used differs from the current widely used Dropbox[127] approach which does use encryption (256-bit AES algorithm) for storing user data, but the data is not encrypted locally and the encryption key of the data stored in the cloud is not owned, nor known by the user, only by Dropbox.

The drawback of the user storing the encryption key is that she needs to be in control of the security of this key, because in case the key is compromised or lost, the data could not be recovered.

Encryption of data in transfer

Protocols:

- **HTTPS[128]**: Hypertext Transfer Protocol Secure (HTTPS) is not a protocol in itself but refers to the use of the Transport Layer Security (TLS) (successor of Secure Sockets Layer protocol - SSL) cryptographic protocol over the ordinary HTTP protocol. This way both, the header and the request/response load of the HTTP messages are encrypted. Connecting to a website using HTTPS allows encrypting most metadata, but attackers can learn on source and destination IP addresses and domain name. HTTPS is also vulnerable to CCA cryptographic attack[128]. Although HTTPS has already been adopted by many popular websites and social networks (Yahoo, Google, Facebook, Twitter, etc.), its widespread use is still in its infancy.
- **IPSec**[129]: Internet Protocol Security (IPsec) is a protocol suite for securing Internet Protocol (IP) communications by authenticating and encrypting each IP packet of a communication session[129]. *IPSEC supports network-level data integrity, data confidentiality, data origin authentication, and replay protection. IPSEC is integrated at the Internet layer (layer 3), it provides security for almost all protocols in the TCP/IP suite[130]. IPSec supports two encryption modes: Transport and Tunnel. Transport mode encrypts only the data portion (payload) of each packet, but leaves the header untouched. The more secure Tunnel mode encrypts both the header and the payload[131].* IPsec has been widely deployed to implement Virtual Private Networks (VPNs).
- **TCP Stealth[132]**: *is a proposed modification of the Transmission Control Protocol (TCP) to hide open ports of some TCP services from the public, in order to impede port scans[133]*. The modification proposal was initiated by Julian Kirsch, Dr. Christian Grothoff, Jacob Appelbaum, and Dr. Holger Kenn after the key discoveries about the details of the HACIENDA[134] surveillance

[120] https://spideroak.com/
[121] https://www.wuala.com/
[122] https://www.boxcryptor.com/
[123] http://www.cloudfogger.com/en/
[124] http://seafile.com/en/home/
[125] http://sparkleshare.org/
[126] https://pyd.io/
[127] https://www.dropbox.com/
[128] http://en.wikipedia.org/wiki/HTTP_Secure
[129] http://en.wikipedia.org/wiki/IPsec
[130] http://secure.vpn4all.com/knowledgebase.php?action=displayarticle&id=646
[131] http://www.cisco.com/c/en/us/td/docs/net_mgmt/vpn_solutions_center/2-0/ip_security/provisioning/guide/IPsecPG1.html
[132] https://datatracker.ietf.org/doc/draft-kirsch-ietf-tcp-stealth/
[133] http://en.wikipedia.org/wiki/TCP_Stealth
[134] http://www.heise.de/ct/artikel/NSA-GCHQ-The-HACIENDA-Program-for-Internet-Colonization-2292681.html

program, which **consists in scanning the entire IP address space of a country for vulnerabilities that can be exploited** As of August 2014 the RFC is an IETF Internet Draft specification.

Tools:

- **HTTPS Everywhere**[135]: This software is the result of a collaboration between The Tor Project and the Electronic Frontier Foundation. HTTPS Everywhere is an add-on for Mozilla Firefox, Google Chrome, and Opera that encrypts communications with many major websites, allowing private browsing. The software helps in the definition of HTTPS rule-sets to define which domains are redirected to HTTPS and how. The enhancement over HTTPS is marketed as: "*Many sites on the web offer some limited support for encryption over HTTPS, but make it difficult to use. For instance, they may default to unencrypted HTTP, or fill encrypted pages with links that go back to the unencrypted site. The HTTPS Everywhere extension fixes these problems by using a clever technology to rewrite requests to these sites to HTTPS*." Note that not all websites are able to work with HTTPS Everywhere, although an increasing number does.

Protection for email services

Protocols:

- **Prism-proof email (PPE) protocol**[136]: PPE by IETF enables encrypted email conversations based on PKI. The preferred cipher set being RSA-2048, AES-256 and SHA-2-512. PPE *puts every individual in charge of their own personal PKI hierarchy. The root of this hierarchy is a personal master key that does not expire until it is either replaced by a new root key or the holder dies.* Therefore, the user of PPE does not need to rely on Trusted Third Parties. Currently PPE does only support Live Mail, but support for more webmails is expected in the future.
- **Bitmessage**[137]: is a protocol for trustless decentralized peer-to-peer encrypted communications. *It encrypts messages, masks the sender and receiver of messages from others, and guarantees that the sender of a message cannot be spoofed, without relying on trust and without burdening the user with the details of key management*[138].

Tools:

- **Sendinc**[139]: Sendinc is free a web-based encryption email service for end-to-end encryption. Sendinc uses 256-bit SSL encryption code and works with any email client, and from any web-enabled device.
- **Enigmail**[140]: Enigmail is a data encryption and decryption extension for Mozilla Thunderbird and the SeaMonkey internet suite. It enables users to write and receive email messages signed and encrypted with the OpenPGP standard. Enigmail works under Microsoft Windows, Unix-like, and Mac OS X operating systems.
- **Mailvelope**[141]: Mailvelope is a browser add-on for Google Chrome and Firefox that offers encrypted email communication based on OpenPGP. It can be configured to work with nearly any web mail provider, e.g. Gmail, Yahoo, Outlook and GMX.

[135] https://www.eff.org/https-everywhere
[136] http://prismproof.org/index.html
[137] https://bitmessage.org/wiki/Main_Page
[138] https://bitmessage.org/bitmessage.pdf
[139] https://www.sendinc.com/
[140] https://enigmail.net/home/index.php
[141] https://www.mailvelope.com/

- **Gnu Privacy Guard**[142]: *GnuPG is a complete and free implementation of the OpenPGP standard as defined by RFC4880 (also known as PGP).* GnuPG allows encrypting email conversations provided both ends use it, similarly to OTR protocol.
- **StartMail**[143]: StartMail offers easy-to-use PGP encryption and TLS based secure email exchange.
- **GPG Tools**[144]: GPG Suite is a free software tool suite that includes GPG Mail and GPG key management among other features. GPG Mail is an open source plugin for Apple's built-in Mail program that allows encryption features (encrypt, decrypt, sign and verify mails) based on OpenPGP.

Protection for Web browsing

- **TOR**[145]: TOR is a free software for Windows, Mac OS X, Linux/Unix, and Android together with an open network that helps protecting the confidentiality of the communications by making it difficult to do traffic analysis. *Individuals use Tor to keep websites from tracking them and their family members, or to connect to news sites, instant messaging services, or the like when these are blocked by their local Internet providers*.

  Tor establishes a network of virtual tunnels (encrypted connections) between source and destination that is incrementally built. *Instead of taking a direct route from source to destination, data packets on the Tor network take a random pathway through several relays that cover your tracks so no observer at any single point can tell where the data came from or where it's going. Tor only works for TCP streams and can be used by any application with SOCKS support*.

  There are a significant number of tools that do exploit Tor system capabilities and network, for example **Tails**[146] free open-source operating system, and other tools explained in the sections below.

- **Tunnelling or VPN services:** Technology to create Virtual Private Network (VPN) connections allows private networks to be extended over public networks, such as the Internet. The VPN services prevent user IP-addresses from being visible to third parties. There are multitude of VPN service providers and prices. Detailed information on collections and ranking of such services is also available in the Web (see for example [147] and [148]). The risk is that **not all VPN services used today are as anonymous as they claim**; some of them do log network traffic which might reveal users IP-addresses and other data. Torrentfreak reports on the logging policies of VPN services[149]. Besides, in some cases the commitment to fulfil the claimed logging policy can be called into question and there is no guarantee that the anonymous VPN service is not handling the user data over third parties. Particularly, VPN services headquartered in countries with doubtful democratic regimes are among those suspicious to share user data with third parties.

- **Do Not Track**[150]: The Tracking Preference Expression or commonly known "Do Not Track" (DNT) is a working draft standard by W3C. DNT is a HTTP mechanism that allows the

[142] https://www.gnupg.org/
[143] https://live.startmail.com/
[144] https://gpgtools.org/
[145] https://www.torproject.org/
[146] https://tails.boum.org/
[147] http://www.top10bestvpn.com/?kw=anonymous%20vpn&c=49593689771&t=search&p=&m=e&adpos=1t2&Dev=c&devmod=&mobval=0&a=1031&gclid=CJifhKacjMICFXDItAodT3EAXQ
[148] http://www.abine.com/blog/2013/vpn-that-protects-your-privacy/
[149] https://torrentfreak.com/which-vpn-services-take-your-anonymity-seriously-2014-edition-140315/
[150] http://www.w3.org/TR/tracking-dnt/

website visitors to inform the websites about their preferences regarding tracking. Most main web browsers (Mozilla Firefox, Safari, Internet Explorer, Chrome and Opera) have the mechanism in place and the user only needs to turn it on to enable the allow/do not allow tracking signal be sent in the HTTP headers. Nevertheless, due to its voluntary character, most websites do not honour the method and disregard the signal or they do honour it but misinterpret the preferences[151]. In consequence, **the method has so far been unsuccessful**.

- **Blur[152]**: Formerly named DoNotTrackMe, Blur is a product of Abine, a Boston based privacy company that among other features (such as encrypted passwords, masked emails, masked cards and auto-filling) supports tracker blocking on web browsing activity. Blur is available in add-on versions for the Microsoft Internet Explorer, Mozilla Firefox, Google Chrome, and Apple Safari browsers. It offers also integrated mobile experience for Android and iPhone. Tracking blocker is included in the free version.
- **Disconnect**[153]: Disconnect is an open-source software that lets the user visualize and block the websites that invisibly track the user personal information. Disconnect is available for Chrome, Firefox, Safari and Opera web browsers and the mobile version as DisconnectMobile for Android. Disconnect needs permission to create a virtual private network (VPN) connection to enable the user disabling tracking by third parties like Facebook, Google, Twitter, and Yahoo, and lets him/her mask the IP address and the location of the VPN servers, in order to browse privately. **The program can also anonymize search queries** in the search engine of choice by blocking identifying cookies *not just changing the appearance of results pages, while staying logged into other services*[154].
- **NoScript**[155]: NoScript is a free and open source add-on for Firefox, SeaMonster and other Mozilla-based browsers. The add-on gives the user the power to specify the sites they trust and only those trusted websites will be allowed to execute active content like JavaScript, Java, Flash and other plugins. NoScript is therefore **a powerful protection against cross-site scripting attacks and clickjacking a**ttacks.
- **Orweb**[156]: Orweb is a free private browser for Android built by The Guardian Project and approved by the Electronic Frontier Foundation. It uses Tor to evade tracking and circumvent network restrictions. The application offers a flexible cookie control, so the user can hide which device is been used and block Flash code from being executed. Orweb requires the free Orbot plugin to be installed too.
- **Onion Browser**[157]: This application allows iOS users to access the Tor network, or browse the web by encrypting their comsmunications and hiding their IP addresses.

Protection for Chat

- **TorChat** *is a decentralized anonymous instant messenger that uses Tor hidden services as its underlying Network, in other words it communicates over the Tor network through the .onion URL protocol*[158]. The use of **TorChat provides end-to-end encryption for secure text messaging and file transfers**. TorChat versions run in Windows, Linux and both iPhone and Android smart

[151] http://www.computerworld.com.au/article/545777/_do_track_oh_what_heck_go_ahead/
[152] https://dnt.abine.com
[153] https://disconnect.me/
[154] http://infosecaffairs.blogspot.in/2014/09/toolsfor-internet-counter-surveillance.html
[155] http://noscript.net/
[156] https://play.google.com/store/apps/details?id=info.guardianproject.browser&hl=en
[157] https://itunes.apple.com/us/app/onion-browser/id519296448?mt=8
[158] http://www.deepdotweb.com/jolly-rogers-security-guide-for-beginners/tor-chat/

phones. The TorChat for Mac users is still in beta version. The program is free open source and is available in GitHub[159].

- **Off-the-Record Messaging**[160] (OTR) protocol serves to encrypt instant messaging conversations over the instant-messaging service of choice, such as gChat or AIM. Both ends in the conversation need to install a software client that supports OTR protocol, first create the keys and then verify each other's encryption keys (in order to avoid man in the middle attacks[161]). OTR uses a combination of the Diffie–Hellman key exchange with 1536 bits group size, the AES 128 and the SHA-1 hash function. *And unlike PGP, which can be a little daunting to learn and use securely,* ***OTR is quite easy to setup and use and provides a pretty good user experience***[162]*. Using* ***OTR only encrypts the contents of your chat conversations but not the metadata related to them****. This metadata includes who you talk to and when and how often you talk to them.*

  "*With the OTR protocol, when both participants in a conversation agree to start an OTR session, the clients set up an encrypted channel with Diffie-Hellman key exchange, then perform a mutual authentication routine inside that channel to verify each other's identity. After the setup, a new key exchange is performed on every message sent, based on incrementing the previously acknowledged key. The participants can independently verify each other's identity using the "Socialist Millionaires' Protocol" (SMP) which allows mutual verification without exchanging private data*"[163].

  Some of the implementations include:

  - **Pidgin**[164]: is free and open source software that runs on Windows, Linux, and other UNIX operating systems and supports multiple chat networks such as AIM, ICQ, Google Talk, Jabber/XMPP, MSN Messenger, Yahoo!, Bonjour, Gadu-Gadu, IRC, Novell GroupWise Messenger, Lotus Sametime, SILC, SIMPLE, MXit, and Zephyr.
  - **Adium**[165]: is free and open source software for Mac OS X that supports AIM, MSN, XMPP (Jabber), Yahoo, and more.
  - **MirOTR**[166]: is a plugin for allowing OTR over Miranda[167] open-source Instant Messaging service.
  - **Cryptocat**[168]**:** This popular and easy-to-use open source software allows instant set up of encrypted chats based on OTR protocol. Cryptocat versions run in Chrome, Firefox, Safari, Opera, OS X and iPhone platforms.
  - **Jitsi**[169]: a free and open source multiplatform for VoIP, videoconferencing and instant messaging that runs with Windows, Linux and Mac OS X. It supports AIM, ICQ, Google Talk, MSN Messenger, Yahoo!, SIP, and XMPP (Jabber).
  - An **OTR library in C#** licensed under The Code Project Open License (CPOL) is available for Windows, Linux and Android platforms[170].

[159] https://github.com/prof7bit/TorChat
[160] https://otr.cypherpunks.ca/
[161] https://freedom.press/encryption-works#otr
[162] http://www.bitcoinnotbombs.com/beginners-guide-to-off-the-record-messaging/
[163] http://www.linux.com/learn/tutorials/341904-weekend-project-secure-instant-messaging-with-off-the-record
[164] https://pidgin.im/
[165] https://www.adium.im/
[166] http://code.google.com/p/mirotr/
[167] http://www.miranda-im.org/
[168] https://crypto.cat/
[169] https://jitsi.org/
[170] http://www.codeproject.com/Articles/644318/Off-The-Record-OTR-Security-Protocol

Protection for Web searches

- **DuckDuckGo**[171]: DuckDuckGo is a popular search engine that does not collect personal information of its users, and therefore all the users are served the same search results for a given search term.
- **Ixquick**[172]: Ixquick is a metasearch engine that encrypts all searches and claims that it does not record the user's IP address, nor share any personal user information with any third party search engine or with the provider of its sponsored results[173]. ***It does, however, collect and store "limited" user information*** *– the date and time of a search, as well as the browser and platform used for the search*[174].
- **Startpage**[175]: Startpage is a search page offered by Ixquick that provides Google search results combined with the privacy policy of Ixquick. *Like DuckDuckGo, Startpage doesn't use cookies, it immediately discards IP addresses, and it doesn't keep a record of searches performed*[176]. Startpage removes from each search query all identifiable information about the searcher, and submits the query to Google search engine. This way, Google cannot trace back the query to the searcher, only to Startpage servers.
- **Blekko**[177]: Blekko search engine is also privacy-conscious. ***Blekko does log personally identifiable information, but deletes it within 48 hours.*** *In contrast, Google stores this information for 9 months – and then anonymizes it without actually deleting it*.
- **Ask**[178]: The Ask search engine provides a feature setting AskEraser which allows the user to indicate search preferences and enforce that cookies are deleted and only a single cookie is left (indicating that AskEraser is enabled). When AskEraser is enabled, the search activity will be deleted from Ask.com servers (**not from third parties**). **The search history will be logged if a critical error occurs** (until recovery) **or under a law enforcement request**.

Privacy aware Operating System

It is recommended that users install security and privacy aware operating systems (OS) on their devices, in order to prevent malware to be installed and/or executed, prevent attackers from accessing the software running on the device, facilitate the creation of VPNs and tunnelling over untrusted connections and support a number of other security features. In the last years some robust implementations have appeared. Two of the most popular are:

- **Qubes OS**[179]: *Qubes OS is an open-source operating system designed to provide strong security for desktop computing using Security by Compartmentalization approach. Qubes OS is based on Xen, the X Window System, and Linux, and can run most Linux applications and utilize most of the Linux drivers*. Qubes OS follows a Security by Isolation approach by facilitating the creation of multiple security domains implemented as lightweight Virtual Machines (VMs) running under the Xen hypervisor. The implementation provides strong isolation among these domains so that an attacker who manages to compromise a domain cannot access all the software running in the other domains. From Version 2 of the Qubes OS, Microsoft Windows virtual machines can be set up that allow running native Windows applications.

[171] https://duckduckgo.com/
[172] https://ixquick.com/
[173] https://www.ixquick.com/eng/privacy-policy.html?
[174] http://searchengineland.com/scroogles-gone-heres-who-still-offers-private-searching-112275
[175] https://startpage.com/
[176] http://www.howtogeek.com/113513/5-alternative-search-engines-that-respect-your-privacy/
[177] http://blekko.com/
[178] http://www.ask.com
[179] https://qubes-os.org/

- **OpenBSD**180: OpenBSD is a free, open-source multi-platform 4.4 Berkeley Software Distribution (BSD)-based UNIX-like operating system. Proactive security and cryptography are two of the features highlighted in the product together with portability, standardisation and correctness. Its built-in cryptography and the pf packet filter make OpenBSD suitable *for use in the security industry, for example on firewalls, intrusion-detection systems and VPN gateways*[181].
- **TAILS**[182] Tails is a complete operating system, designed to be used from a DVD, USB stick, or SD card that, by relying on the Tor anonymity network, aims to preserve privacy and anonymity. Tails comes with several built-in applications pre-configured with security in mind.

[180] http://www.openbsd.org/
[181] http://en.wikipedia.org/wiki/Openbsd
[182] https://tails.boum.org/about/index.en.html

2 THEME 2 - CRYPTOGRAPHY RELIABILITY IN A "POST SNOWDEN" WORLD

2.1 Question 7

What are the known "exploits" that can be used to compromise the use of current encryption in the Internet for mass surveillance purposes, and what can end-users do to better protect themselves?
What are examples of real-world cases where supposedly secure cryptographic/encryption systems have been compromised and have led to unintended disclosure of information?

The new generation of encryption technology is proficient enough to avoid deterministic brute force attacks if its implementation and configuration parameters are well set up. Major attacks occur when implementations of current encryption technologies do not faithfully comply with their specifications, or when bugs and flaws – sometimes deliberately - arise at code level[183],[184].

However, mass surveillance attacks are not only confined to a specific encryption factor. There are many other means (social engineering, malware, Trojans, worms) of exploiting IT-systems, as the following examples of exploits compromising encryption and other security mechanisms show:

"goto fail" SSL vulnerability

Secure Socket Layer (SSL) is a protocol utilized to encrypt and authenticate against server connections: i.e. when connecting for consulting bank data, the information needs to be encrypted and the connection needs to be confirmed, so as to avoid connecting to a fraudulent web site. For exchanging information in a secure way, the web browser and the server have to share a key. In this case, the server must send a series of initial parameters that must be digitally signed. This signature has to be verified by the client (navigator).

In some proprietary implementations (such as Apples iOS) for verifying this piece of code (signature) a 'goto fail' instruction was (either deliberately or not) duplicated in the source of the verification code (see Figure 4), which led to a severe vulnerability of the SSL protocol:

"If.. then" clauses keep updating the hash; if it fails then the line "goto fail" is executed. Since there is a duplicated "goto fail" (in line 12), it will always be executed. This means that the program flow will always pass through this instruction and therefore return a 0 that indicates that everything was executed correctly. Coincidentally, this second "goto fail" instruction is executed just after the clause verifying the signature and thus jumps over this verification. The returned value will always be a 0 and any application connected to it will reason that the signature is correct.

This flaw would allow anyone to eavesdrop on the secured connections by applying a man in the middle attack. A back server would interfere between the bank account on the server side and the browser on the client side, enabling an attacker to get access to the data exchange.

[183] https://gotofail.com/
[184] http://heartbleed.com/

```
3   static OSStatus
4   SSLVerifySignedServerKeyExchange(SSLContext *ctx, bool isRsa, SSLBuffer signedParams,
5                                    uint8_t *signature, UInt16 signatureLen) {
6       /* ... */
7
8       if ((err = SSLHashSHA1.update(&hashCtx, &serverRandom)) != 0)
9           goto fail;
10      if ((err = SSLHashSHA1.update(&hashCtx, &signedParams)) != 0)
11          goto fail;
12          goto fail;
13      if ((err = SSLHashSHA1.final(&hashCtx, &hashOut)) != 0)
14          goto fail;
15
16      /* ... */
17
18  fail:
19      SSLFreeBuffer(&signedHashes);
20      SSLFreeBuffer(&hashCtx);
21      return err;
22  }
```

Figure 4: Duplicated 'got fail' statement in Apple's SSL verification code

The "goto fail" flaw affected Apple mobile products that run iOS 6 and iOS 7 as well as desktop products that run OS X 10.9 (Maverick). iOS 6 was launched in September 2012, which means this vulnerability existed about 2 years before being detected and patched.

Heartbleed

Heartbleed is a security bug revealed in April 2014. It appeared in an OpenSSL cryptography library which is commonly utilized for implementing TLS (Transport Layer Security). Its vulnerability proceeds from the improper validation on the implementation of the "Heartbeat" extension of TLS. This bug is categorized as a buffer over-read that consents reading more data than should be permitted. At the time of disclosure, around half a million of the Internet's secure web servers certified by trusted authorities were alleged to be vulnerable to the attack, permitting theft of the servers' private keys and users' session cookies and passwords[185].

Shellshock

Beside Heartbleed, Shellshock has been identified as a bug that possibly infects millions of workstations running on Linux and Apple's Mac OS. The bug was discovered in a software component called Bash (Bourn-Again Shell) which is included in many Linux/Unix systems as the default command-line interface. Due to this vulnerability, it is possible to execute random commands and take control of the system utilizing Bash. More than 500 million devices could be infected by this flaw, which seems to have existed since version 1.03 of Bash that was released in September 1989. There are already numerous patches available that (sometimes partially) solve the problem, but not all affected devices and servers have been patched already at the time of writing this document (Oct. 2014).

Phishing attacks: Apple iCLOUD. Celebgate // The Fappening

Recent Phishing attacks that have compromised services provided by Apple's iCLOUD were prominently covered in all types of media. Attackers targeted celebrities and exposed some of their private photos that were saved on the iCloud platform on the Internet. The attackers used Phishing techniques to access these private areas and files on iCloud. Phishing attacks consist in tricking users into revealing their login credentials by using spoofed emails, websites and the like. In the case of "Celebgate", Apple determined and confirmed that the hacks did not breach security measures and were not based on encryption failures, but on social engineering attacks. The victims had received

[185] http://news.netcraft.com/archives/2014/04/08/half-a-million-widely-trusted-websites-vulnerable-to-heartbleed-bug.html

legitimate looking company emails requiring private information (personal material, passwords, credit card numbers, etc.) which they had answered. Phishing attacks derive from gathering the username – usually an email address associated with the account. In order to crack passwords of celebrities' accounts, hackers used a program (intelligent brute force) that cracked the stars' secret questions by making repeated guesses using personal information available online. After the Celebgate phishing attack other exploits such as the Kelihos (Waledac) botnet appeared for exploiting the iCloud system and capable of sending (spam) emails to specific Apple customers, recommending urgently to check their AppleID.

McAfee, an antivirus vendor, tested end users' skills for identifying phishing emails and showed that 80% of 16.000 users fell for at least one in seven phishing emails. Currently antimalware companies do provide protection against phishing attacks (Symantec, Kaspersky, Trend, Panda, etc.), but many final users do not adequately protect their computer by updating the latest patches of Operating Systems or applications. Also, suspicious requests on personal information are generally answered without applying a minimum of precaution (and/or common sense) and thus handing over personal and intimate data to potential attackers.

Botnets

Botnets are networks of devices infected with malware that are remotely controlled by one or more attackers and used to perform distributed tasks on behalf of the "owners" of such botnets. Botnets are primarily used for spamming or supporting distributed denial of service attacks. The infection of a computer with the malware, required for forming part of a botnet, is commonly achieved through drive-by infections (a malicious website downloads and installs the malware on the computer when the user visits it) or by emails with infected attachments. Some botnets, like for instance "BredoLab" and "Mariposa" (both have been dismantled a few years ago) have reached an extension of tens of millions of infected PCs and compromised personal, corporate, government and university computers in more than 190 countries including account information, usernames, passwords, banking credentials, and credit card data. Security companies such as Afferent Security Labs, Symantec, Trend Micro, FireEye, Umbra Data and Damballa have announced offerings to counter botnets.

APT attack : advanced persistent threats. Industrial field: dragonfly

Dragonfly is another attempt to attack industry and particularly the electricity sector with cyberwar tools. Eastern European countries and Russia collaborated in developing a sophisticated attack on electrical companies based in NATO countries[186]. In this case, the objective was not the espionage itself but the sabotage: service disruption. The Dragonfly group uses attack methods which are centred on extracting and uploading stolen data, installing further malware onto systems and running executable files on infected computers. Since industrial control systems are usually connected to Internet (see SHODAN search tool) they are being attacked by applying complex cyberwar strategies.

Dragonfly is a compendium of attacks which alternates phishing techniques with the usage of RAT (Remote Access Tools) based backdoors such as Oldrea (also known as Havex or Energetic Bear). Oldrea extracts an infected computers' information once it is installed and sends this information in an encrypted way to a server controlled by attackers. Then by adapting and using the "Karagany" Trojan the attackers are able to run additional plugins, such as tools for collecting passwords, taking screenshots, and cataloguing documents on infected computers. Finally, through "trojanizing" three different Industrial Control System equipments, the attacker could eventually end up with the objective of sabotaging a power grid. Therefore, Dragonfly is a complex attack that employs multiple attack vectors: phishing, exploit kits, remote access tools and trojanized software.

[186] http://www.symantec.com/connect/blogs/dragonfly-western-energy-companies-under-sabotage-threat

Zero-day vulnerabilities

A zero day vulnerability is a defect in software that is unknown to the vendor and that is subject to potential exploitation until the vendor becomes aware and fixes it. Vendors usually release a patch or software update in order to fix security flaws, once they are detected and published. Discussions on zero day vulnerabilities can be found in specialized and private security forums and web pages, but there does also exists a "black/grey market" in which zero day vulnerabilities are sold[187]. It is suspected that national security agencies, as well as criminal organisations make use of this offer, in order to devise new exploits for which no known remedy exists[188].

It is impossible for a company to totally protect against zero day attacks. However, there are a number of recommended steps that help to reduce the impact of a potential attack:

1. Prevention

Companies should apply security best practices which include the installation of firewalls and other perimeter security devices and keeping operative systems, applications and particularly antivirus tools updated at any time. Furthermore, it is a good practice to regularly execute security audits on the systems, in order to detect software vulnerabilities as early as possible.

2. Real time monitoring and protection

A number of tools like Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), security monitoring tools, etc. are available that either prevent from, or alert of malware infection. The deployment and usage of these tools helps in reducing the impact of malware and virus infections by allowing to devise early solutions in form of patches and updates.

3. Planned incident response

When a company is affected by a zero day attack it is crucial to have implemented appropriate incident response procedures, including roles and responsibilities, to minimize damage and business disruption.

4. Preventing the spread

Spread prevention basically consists in isolating networks and opening only those that are necessary for the continuity of the business.

Brute force attacks

Brute force attacks[189] are well known attacks that explore all or part of possible values in order to extract the password of an end user. The strength of a brute force attack is inversely proportional to the complexity of employed encryption key. Indeed, brute force is not dealing with cryptography vulnerabilities but with key vulnerabilities. The simpler the key (for instance, by generating passwords of 4 numbers), the easier it is to extract the key by this mechanism.

Today brute force attacks are often used and quite efficient when preceded by a social engineering attack which reveals part (or all of) of the encryption key, or provides knowledge that helps to reduce the possible values for extracting the target's key (*Celebgate* is a clear example for this.)

Safeguards

Different protection mechanisms and behaviours could be used as antagonizing measures for 4 cryptography problems:

[187] https://www.mitnicksecurity.com/shopping/absolute-zero-day-exploit-exchange
[188] http://insidecybersecurity.com/Cyber-General/Cyber-Public-Content/policy-debate-looms-on-us-role-in-market-for-zero-day-cyber-threats/menu-id-1089.html
[189] https://www.owasp.org/index.php/Brute_force_attack

1. Obtaining "encryption keys" through adequate intelligence operation on one of the two end-points.
    - Best practices are oriented to generate strong keys (both symmetric and asymmetric ones). These keys should be:
        - Long (>8 characters) and generated by using a mix of alphanumeric and special symbols.
        - Random numbers
        - Dynamic (i.e. keys should be renewed periodically)
    - PGP tools[190] make this possible by using a public/private encryption infrastructure. PGP's public/private key generator allows creating strong encryption keys that ensure high confidentiality and integrity factors.
    - One of the most effective and easy to apply practice is that of generating passwords with a sufficient length. The table below shows the relation between the length of a password (i.e. number of characters) and the effort needed to break it.

Number of characters in Password	Average days needed to crack password
10	5
11	306
12	18.976,5
13	1.176.549
14	72.946.048

Table 2: Relation between password length and effort for cracking it (source theWire[191])

2. Exploiting security vulnerabilities (backdoors, side channel attacks, bugs, viruses, APT, BotNets, etc..) in the set-up of one or both of the "end-points".
    - This problem is addressed by *security by default* and *security by design* paradigms. The first one refers to the most secure use of default configuration when using software programs (which is not necessarily the most user friendly, or performance effective one, but often the best option to avoid backdoors); the second refers to the engineering phase of a particular software product and the consideration of security requirements from its early stages.
    - Both paradigms help to avoid vulnerabilities in the engineering and operation phases, but need a validation or certification scheme (above all at engineering phase) which guarantees that IT products do not contain any known security vulnerability.
3. Exploiting weaknesses (bugs, software design flaws and backdoors) in the encryption programme to enable decrypting the information without the need of possessing the original encryption key(s) from one of the end-points.
    - At this stage, security vendors for SSL/TLS and IPSec and other types of encryption programmes need to be validated and certified for avoiding:
        - Unfixed SSL/TSL and IPsec protocol vulnerabilities in the different versions of the implementation of the protocols
        - Interoperability problems caused by different implementations (from a functional and non-functional point of view)

[190] For instance, https://www.igolder.com/pgp/encryption/
[191] http://www.thewire.com/technology/2012/08/advice-hacker-password-security-best-practices/56343/

2.2 Question 8

What types of backdoors exist in commercial IT products?
How can end-users possibly detect and disable them?
What are the most and less secure products or family of IT products on the market?
How can the end-user distinguish between "backdoor free" IT products and "non-backdoor free" IT products at purchase time and during their whole life cycle?

IT products address a great range of tools, platforms, software components and systems worldwide. There also exists a large variety of possible backdoors for the whole range of IT products. As Professor Matt Blaze[192] points out, building secure software is a hard enough challenge in its own and therefore attempting to include security in the whole range of (interacting) products is little more than wishful thinking. Indeed, most IT products were not conceived and engineered by "security by design" principles and are neither operated by "security-by-default" doctrines. However, this is gradually changing towards a more security-formalized domain.

Commercial IT products that are likely to be infected could be categorized according to the following scheme (another dimension for structuring IT Products is to categorize on terms of Open Source[193] and Proprietary Software):

- **Devices:** commonly ad-hoc firmware and software for particular purposes. Normally security features are not included as default requirements. A security breach is relatively easy to reach if the attacker domains low level programming, reverse engineering and attacking techniques.
- **Data centre systems:** a data centre is a computer system and associated components (telecommunications and storage subsystems). Data centres are constructed for enabling business continuity and therefore their main requirement stands on availability. With the increasing uptake of cloud computing, security oriented challenges such as Cloud Data Centre geographical legal framework, legal service level agreements or privacy aspects are being addressed that go beyond availability aspects.
- **Enterprise software:** is software utilized to satisfy organizations (businesses, schools, user groups, retailers or governments) requirements. It is the integral part of information systems and it also encompasses web based software production. Services provided by enterprise software are typically business-oriented tools such as online shopping and online payment processing, interactive product catalogue, automated billing systems, security, enterprise content management, IT service management, customer relationship management, enterprise resource planning, business intelligence, project management, collaboration, human resource management, manufacturing, enterprise application integration, and enterprise forms automation. Security by itself is being inserted as a non-functional and mandatory requirement into this type of IT products, making emphasis on the confidentiality feature of the digital assets and focussing on authentication techniques.
- **IT Operation and Management**: ITOM vendors (main are BMC Software, CA Technologies, IBM and HP) provide a way to organize and synchronize the use of IT tools to provide better value to the business as well as transformational services to ensure balance between people, process and technology. These tools are principally marketed with a focus on processes rather than technology, which results in a lesser attention to security related aspects than in other product categories.

[192] Matt Blaze is directing the Distributed Systems Lab University of Pennsylvania
[193] http://timreview.ca/article/157

- **Telecom services:** we refer to the software oriented to support the provision of telecommunication services (by provider). Much effort has been devoted to security and privacy issues for telecom services without much success, since government security policies concerning encryption capacity prevent operators from implementing full-fledged end-to-end encrypted communication channels.
- **Industrial software:** software with long term amortization periods that focus on the functionality and availability (e.g. SCADA systems) and usually not on security and privacy. Many vulnerabilities arise because no proper patching policy is adopted. Security breaches are imminent in this field (a well-known example of this is the infection of Iranian nuclear facilities with the Stuxnet worm).

A backdoor is a hidden entrance to a computer system that can be used to bypass security policies[194]. There are two main types of backdoors[195]:

- **Conventional:** There are different types of backdoors affecting conventional commercial IT Products. The main types are:
    - Exposition of administration and/or management interfaces (normally through brute force attack)
    - redundant interfaces/functions/features (due to faulty procedures or by mistake)
    - hidden parameters (backdoors by design)
    - redundant users (i.e. default users such as guest)
    - authorisation for 3rd party access (again backdoors by design).
- **Unconventional:** Unconventional backdoors are posing a high risk since they exploit system weaknesses and zero-day vulnerabilities to deploy very difficult to detect ad-hoc system attacks. The most frequent types of unconventional backdoors are the following:
    - authentication and authorization between application components (application level backdoor vulnerability)
    - old users in systems (those originally created according to standards procedures and not being aware of future security needs)
    - flawed hardening (enabling multiple attack vectors – OS, Application Servers and Database hardening)
    - exposed configuration data (weaved into systems' architecture)
    - lack of separation between environments (security vulnerability leading to a backdoor)

Mobile backdoors

"*Upon receipt of a valid search warrant, Apple can extract certain categories of active data from passcode locked iOS devices. Specifically, the user generated active files on an iOS device that are contained in Apple's native apps and for which the data is not encrypted using the passcode ("user generated active files"), can be extracted and provided to law enforcement on external media. Apple can perform this data extraction process on iOS devices running iOS 4 or more recent versions of iOS. Please note the only categories of user generated active files that can be provided to law enforcement, pursuant to a valid search warrant, are: SMS, photos, videos, contacts, audio recording, and call history. Apple cannot provide: email, calendar entries, or any third-party App data*."[196]

In iOS 7 the situation varies since encrypted data is not mapped to the password but rather encrypted through a key stored in a secure element. Library and caches are free of encryption rules. Therefore, after the reboot and unlock tasks, most of encrypted data could be accessed (but it must be unencrypted.)

There are some forensic tools designed for unencrypting accessible encrypted data. Indeed: "*Federal agencies have always been interested in black bag techniques (compromised docking stations, alarm clocks, etc).*"

[194] Microsoft definition
[195] OWASP (Open Web Application Security Project)
[196] https://pentest.com/ios_backdoors_attack_points_surveillance_mechanisms.pdf

Der Spiegel states based on documents of Snowden*: "The documents state that it is possible for the NSA to tap most sensitive data held on these smart phones, including contact lists SMS traffic, notes and location information about where a user has been. In the internal documents, experts boast about successful access to iPhone data in instances where the NSA is able to infiltrate the computer a person uses to sync their iPhone. Mini-programs, so-called "scripts," then enable additional access to at least 38 iPhone features." Actually,* DROPOUTJEEP is a NSA developed a progamme that permits access to iPHONE features. The question is if Apple knows about this fact. Jacob Appelbaum, an independent researcher from the TOR project stated in a conference about this issue[197]: "[the NSA] *literally claim that anytime they target an iOS device that it will succeed for implantation. Either they have a huge collection of exploits that work against Apple products, meaning that they are hoarding information about critical systems that American companies produce and sabotaging them, or Apple sabotaged it themselves.*"

The German magazine "Der Spiegel" reported on a catalogue of more than 50 backdoors the NSA has apparently utilized without the consent and knowledge of concerned IT Product manufacturers. This catalogue was developed by NSAs TAO[198] hacking unit.

Backdoors have also threatened Android systems. Actually an Android backdoor was discovered[199] in Samsung Galaxy (at least nine different models of smartphones and tablets) from which an attacker could remotely gain access to the cited smartphone and steal files and location data or surreptitiously activate a microphone or camera. Indeed, the Replicant Project[200] goes further by saying that "*We discovered that the proprietary program running on the applications processor in charge of handling the communication protocol with the modem actually implements a backdoor that lets the modem perform remote file I/O operations on the file system*" (Paul Kocialkowski in a blog post on the Free Software Foundation site.[201])

The Replicant Project is a free Android distribution running on several devices, a free software mobile operating system putting the emphasis on freedom and privacy/security. They state that a mobile device respecting the users' freedom would/should have:

- Free hardware: should be referred to define systematic processes for producing PCBs (Printed Circuit Boards) with schematics, as a first approach for creating new low cost hardware versions. This is something that does not yet exist.
- Free firmware: generally firmware running within integrated circuits is proprietary. Even for firmware oriented to particular free hardware such as Arduino[202] it is complicated to replace, version or alter its code.
- Free modem system: It is always proprietary[203]. Only one free software GSM stack, called OsmocomBB, exists which however only works for old feature phones.
- Free bootROM and bootloader: bootROM is always proprietary and stored in read only format. The bootloader could be modified as some of them are free, but most of them are cryptographically signed (asymmetrically) which avoids its modification.

[197] http://www.zerohedge.com/news/2013-12-30/how-nsa-hacks-your-iphone-presenting-dropout-jeep
[198] http://www.spiegel.de/international/world/the-nsa-uses-powerful-toolbox-in-effort-to-spy-on-global-networks-a-940969.html
[199] http://www.darkreading.com/mobile-security/samsung-galaxy-security-alert-android-backdoor-discovered/d/d-id/1127675
[200] http://www.replicant.us/
[201] https://www.fsf.org/blogs/community/replicant-developers-find-and-close-samsung-galaxy-backdoor
[202] http://www.arduino.cc/
[203] Based on Replicant Project: "A device with bad modem isolation would allow the modem to access and control key parts of the hardware, such as the RAM, storage, GPS, camera, user I/O and microphone. This situation is terrible for privacy/security as it provides plenty of ways to efficiently spy on the user, triggered remotely over the mobile telephony network"

- Free systems and applications: many well-known operating systems are free (Android[204], Firefox OS[205], Ubuntu Touch[206], Tizen[207]). However the user-space hardware abstraction layers are for the most part proprietary , as well as other software components inside these systems: This means that the community cannot entirely control the source code and confirm the absence of backdoors, a situation that impedes the fulfillment of the Replicant manifesto.

For an end-user it is virtually impossible to distinguish between "backdoor free" and "non-backdoor-free" products, unless the media or pertinent internet fora inform about such back-doors. If a backdoor has been detected, the product manufacturer usually offers a patch or update that fixes the problem. However, it is the user's responsibility to install and apply these security fixes. It is therefore **recommended to configure applications such that they automatically install** (or at least notify the user about) **an available security update**.

2.3 Question 9

In the current generation (2014) of categories of hardware platforms (PCs, Tablets, smartphones) and Operating Systems (Windows, IOs, OSX, Android, Unix, Linux, etc…), are some more prone to contain "malware" allowing to circumvent encryption for mass surveillance purposes?
What are known cases of identified and patched backdoor or malware vulnerabilities (for instance the Chrome design flaws reported in January and April 2014 giving the possibility to an attacker to activate microphones remotely without end-user notification)?

Traditionally, malware has been inoculated on platforms with the highest rates of impact in industry and society. Accordingly, PCs have been vastly infected over decades, with the duo PC-Windows being the platform of highest interest to hackers and attackers.

In the last decade Microsoft has prioritized security aspects through the Trustworthy Computing Initiative[208] and despite the fact that the Microsoft Windows platform continues to be the primary target for attackers, it seems that Microsoft has taken the pulse of the security challenge. But with the advent of massive mobile computing things are changing. The ever growing number of wireless connected (heterogeneous) devices and the resulting societal impact and changing usage patterns makes attackers shift to these mobile platforms, which furthermore can be attacked and potentially damaged by applying new and innovative intrusion approaches.

In the mobile world, Android and iOS are exploited in radically different manners. While Android is free and open source[209], iOS is privative and restricted to Apple's internal policies. However, none of the two operating systems is immune to the possible infection with trojans, bugs or worms.

BlackBerry has long touted its strong encryption features when selling its smartphones to corporate clients. But this protection did also cause problems when the company came under pressure from countries, such as the United Arab Emirates and Saudi Arabia, who wanted Research in Motion (RIM, the manufacturer of BlackBerry phones) to provide access to their customers' secure email (apparently these governments didn´t want citizens to encrypt their communications).

[204] http://www.android.com/
[205] https://www.mozilla.org/es-ES/firefox/os/
[206] http://www.ubuntu.com/phone
[207] https://www.tizen.org/es?langredirect=1
[208] http://www.microsoft.com/en-us/twc/
[209] https://source.android.com/

Linux systems do apparently get less targeted by attackers. But their increased usage in server stacks (e.g. LAMP) together with recently published flaws (see Heartbleed and Shellshock), make this system more and more interesting for hacking attacks.

A novel hardware scenario is the advent of the smartphone which encompasses computer features and resources but still has the voice service which makes it more prone to impact (apart of the mobility and camera factors which make the smartphone the platform that exposes the largest attack surface). Smartphones and their connections are prone to be attacked by other malware from national agencies and groups dedicated to cybercrime. Indeed, the NSA has dedicated much effort to implement tools for exploiting not only mobile and smartphone connections, but also the endpoint. There are three main programmes ongoing:

- TYPHON HX**:** GSM station controlling signals (SMS, calls) and its geo-localization
- CANDYGRAM: similar to TYPHON HX, that permits the remote control though SMS

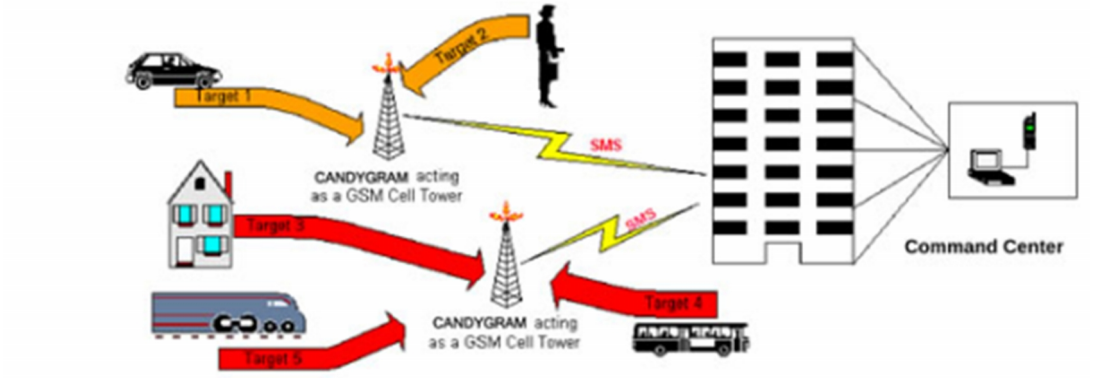


Figure 5: CANDYGRAM Operational Concept

- MONKEYCALENDAR**:** it is a malware for SIM cards that allows remote access to geographical position.

As an example, the iOS7 platform has recently been updated (January 2014) due to a severe vulnerability in its SSL/TLS encryption system. Apple's library was not able to authenticate pair connections and therefore, anybody could listen to apparently secured connections. The same vulnerability was discovered in OS X. There was also a huge coincidence detected between the dates on which such failure appeared and leaks coming from the NSA and the PRISM program.

2.4 Question 10

> ***Is equipment produced in some countries by some manufacturers and solution integrators more prone to include "malware" in the hardware, firmware installed (for backdoors purposes)?***
> ***If yes, what are the causes?***
> ***Is there a risk of "balkanisation of technology" with some countries producing "clean" vs "dirty" products?***

One of the major consequences of the world economic crisis was the reduction of budget, which, in the IT world, led (among other things) to a reduction of security validation and verification, both, in the engineering and operating phases and both, in public and private sectors. The production cost is becoming the most important factor when manufacturing chips and boards and many chip producers are suspect to reduce quality (and security) assurance efforts in order to achieve economies of scale.

This **reduced quality assurance can lead to unintended flaws in the chips design and/or manufacturing process, which on their part can enable (unintended) backdoors**.

Some media reports directly accuse Asian manufacturers of purposefully including backdoors in their chips. The most prominent of these accusations is based on a paper published by Sergei Skorobogatov, a student of Cambridge University, in which he claims to have identified a backdoor in Actel/Microsemi's ProASIC3 chip[210]. However, a subsequent discussion of the findings in the security community showed that the backdoor (a JTAG debugging interface) was already included in the software design of the chip, that was developed by Actel, an American company[211].

In 2011, the US government released the White House Cyber Policy Review[212], warning of risks related to the delocalization of manufacturing plants:

"*The emergence of new centers for manufacturing, design, and research across the globe raises concerns about the potential for easier subversion of computers and networks through subtle hardware or software manipulations. Counterfeit products have created the most visible supply problems, but few documented examples exist of unambiguous, deliberate subversions.*" This suspicion has led Western countries and particularly the Five Eyes coalition to prohibit the usage of certain Chinese built equipment in intelligence agencies[213].

But it's not only the Asian countries that are allegedly designing stealth backdoors into their products. Documents revealed by Edward Snowden affirm that the NSA is secretly and routinely intercepting "*routers, servers, and other computer network devices being exported from the U.S. before they are delivered to the international customers. The agency then implants backdoor surveillance tools, repackages the devices with a factory seal, and sends them on. The NSA thus gains access to entire networks and all their users.*"[214] Other reports suggest that the NSA is also collaborating with major vendors of encryption technology with the objective of gaining access to encrypted information. This is achieved by introducing a flawed generation of random numbers that serve as seeds for encryption in the RSA algorithm[215].

A SPIEGEL report claims that the NSA is using a program out of their ANT toolbox catalogue that "*attacks the firmware in hard drives manufactured by Western Digital, Seagate, Maxtor and Samsung, all of which, with the exception of the latter, are American companies*"[216]

Summarising this information, **it is not possible to pinpoint specific countries or even regions to be more prone than others to include malware in hardware or firmware**. It rather looks as if different parties in all relevant production areas are using such a strategy if national or security interests are concerned.

There are many reasons to include malware in hardware/firmware but two of them are important to mention:

1) Information leakage: Product design espionage
2) Service interruption or malfunction: Unlocking devices, features and gaining access.

[210] https://www.cl.cam.ac.uk/~sps32/ches2012-backdoor.pdf
[211] http://www.cl.cam.ac.uk/~sps32/microsemi_re.pdf
[212] http://www.whitehouse.gov/assets/documents/Cyberspace_Policy_Review_final.pdf
[213] http://securityaffairs.co/wordpress/16748/hacking/spy-agencies-ban-on-lenovo-pcs-due-to-backdoor-vulnerabilities.html
[214] Greenwald, G.: No Place to Hide: Edward Snowden, the NSA, and the U. S. Surveillance State, Metropolitan Books, New York (2014)
[215] http://www.reuters.com/article/2013/12/20/us-usa-security-rsa-idUSBRE9BJ1C220131220
[216] http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994.html

2.5 Question 11

Why are all data communications not systematically encrypted over the Internet?
How to compare the reliability and quality of encryption of email, voice over IP communications (SKYPE, SIP), cloud-computing application services (Dropbox, etc..), email, and generic web traffic today?
Is it all of the same quality and reliability or are there significant differences in the way encryption and cryptography can be broken or circumvented and end-user privacy compromised?

The increasing concern of citizens about their privacy is pushing more and more internet service providers towards offering (communication) services that are secured and encrypted by default. One of the last decisions in this direction was Apple's and Google's announcement[217] of the imminent introduction of fully encrypted (voice and text) communications in their operating systems product family for smartphones. Google also announced in March 2014 that its Gmail service is using "*an always-on HTTPS connection and encrypting all Gmail messages moving internally on its servers*"[218]. Also "Whatsapp", a text-based message application with more than 500 million users worldwide, has switched to encrypting the communication between the phones and their servers[219]. Google's new SPDY gateway[220] is another attempt to perform secured transfers over Internet.

It seems that the main platforms tend to include encryption in webmail using SSL and TLSS (gmail, outlook.com …). Other native applications, such as Outlook using POP3 and SMTP[221], depend on the server policies or client encryption options for a secured utilization (by using for example SMTPs and POP3s).

However, it must be pointed out that these attempts to increase the privacy of end users by enabling strong encryption by default are not welcome by all stakeholders. James B. Comey, the director of the FBI, in a speech published on their website[222] is insistently warning on the danger of the law enforcement "going dark" as a consequence of these encryption efforts. Comey (and others) argues that "*law enforcement needs to be able to access communications and information* [i.e. evidence] *to bring people to justice*, but that with "*sophisticated encryption* [being available to end users for securing their conversations], *there might be no solution, leaving the government at a dead end – all in the name of privacy and network security*". He goes even so far as to call on the U.S. Congress to make unlockable encrypted devices, as they were announced by Apple, unlawful[223].

Various rankings[224] of security vendors' encryption services can be found on the web, but their evaluation criteria are not always transparent. **The quality and reliability of a specific encryption implementation**, or service, **can be compared by evaluating** some **objective indicators**[225], which are:

- Asymmetric cryptography usage for authentication

[217] http://www.bbc.com/news/technology-29276955
[218] http://nakedsecurity.sophos.com/2014/03/21/google-switches-gmail-to-https-only/
[219] https://www.whatsapp.com/faq/en/general/21864047
[220] http://www.google.com/patents/US20130297814
[221] http://www.internode.on.net/support/guides/email/secure_email/
[222] http://www.fbi.gov/news/speeches/going-dark-are-technology-privacy-and-public-safety-on-a-collision-course
[223] http://www.washingtonpost.com/blogs/the-switch/wp/2014/10/17/fbi-director-comey-calls-on-congress-to-stop-unlockable-encryption-good-luck-with-that/
[224] http://encryption-software-review.toptenreviews.com/
[225] See more file encryption system indicators at:
http://en.wikipedia.org/wiki/Comparison_of_disk_encryption_software

- Key strength or Passphrase strengthening
- Cryptography algorithms combination
- Dynamic key re-generation capacity
- Hidden containers
- Pre-boot authentication
- Multiple keys usage for a particular process
- Hardware acceleration in particular situations, explicitly designed for cryptography.

Such an evaluation can be performed in a static way by analysing one by one each of the implementations for the encryption services, not their specification[226]. Different strengths and weaknesses of encryption tools or services can be only be found by analysing their implementation.

Despite the theoretical knowledge that, for instance DES is weaker than 3DES because the latter is reusing 3 times DES, making it 3 times more powerful, or that Blowfish is weaker than AES since its block size continues to be 64 bits, just like DES, it is the quality and correctness of the implemented algorithms that define the quality of a particular encryption.

Software flaws in the implementation of encryption algorithms (e.g. in negotiation, or key exchange functions) **can lead to vulnerabilities** which may be easily exploitable, **regardless of the complexity, or the theoretical strength and quality of the applied encryption technique**.

Since the implementation is the critical part that defines the overall quality of encryption solutions, a process for guaranteeing the satisfaction of these indicators must be applied. This process should cover the design and engineering phase, ensuring that appropriate measures are taken to avoid injecting software errors, but also consider certification of the final product code by external independent laboratories. NIST is providing such a service[227], through which it informed the public about a weakness (that could be exploited as a backdoor) of the Dual_EC_DRBG encryption algorithm and "*recommends that current users of Dual_EC_DRBG transition to one of the three remaining approved algorithms as quickly as possible*"[228]

2.6 Question 12

Are there significant differences of implementation of the Open Internet encryption communication protocols (SSL/TLS and IPsec) across different vendors, platforms and service providers?
Are there interoperability issues between different implementations of the same standard that make encryption for some services and software less reliable than for others?
What is the level of awareness of the public about these possible differences?

Web technology is encrypted through SSL/TLS technology. The protocol itself is correct and is able to maintain confidentiality, integrity and therefore users' privacy. However, concrete encryption implementations are prone to contain bugs.

A protocol is usually defined by its common formal reference specification. This specification can be and is being implemented by different parties, which inevitably leads to different code bases.

[226] http://www.csjournals.com/IJCSC/PDF2-1/Article_22.pdf
[227] http://csrc.nist.gov/groups/STM/cavp/documents/aes/aesval.html
[228] http://www.nist.gov/itl/csd/sp800-90-042114.cfm

OpenSSL[229], for instance, is an open source library often used in Linux platforms (Apache and Nginx) that implements SSL V2, V3 and TLS V1. The "heartbleed" bug[230] detected in (some of) these implementations shows that even open source code which is screened and reviewed by a large community of developers and which had been deployed and operative for years does not guarantee to be free of any software flaw.

Beyond OpenSSL there are more vendors implementing SSL and TLS: Microsoft Windows has its own implementation within its secure channel (Schannel) component[231]. This Schannel implementation was also subject to a vulnerability[232] (for which a patch is available) that would allow spoofing, but seemed to be a mature and proven implementation, until Microsoft published a new "*vulnerability in the Microsoft Secure Channel (Schannel) security package in Windows. The vulnerability could allow remote code execution if an attacker sends specially crafted packets to a Windows server*."[233] Apple, on the other hand, developed its own implementation of the Secure Transport package, which was plagued by the "goto fail" bug[234].

Much of the same problems arise for IPSec[235]. There are several implementations referring to approximately 40 protocols and RFCs (for example key interchange mechanisms[236]). Vendors are also well known such as Microsoft Windows, Solaris, AIX from IBM, Z/OS from IBM and OpenBSD. KAME[237] is also an implementation included in some referential platforms such as MAC OSX, netBSD and FreeBSD.

The most commonly used browsers (Chrome, Firefox, IE, Opera and Safari) support TLS 1.0, 1.1 and 1.2 in their latest versions. However, older versions of these browsers may be prone to vulnerabilities concerning these protocols.

The level of user awareness related to these issues is substantially low and thus, it is important to make the end user understand that implementation errors and bugs are normally fixed by upgrading and updating software and components. This means that **it is the user's own responsibility to keep the applications properly updated**.

2.7 Question 13

> ***What are current practices of encryption for the transfer of personal data between data centres operated by social network service providers such as Facebook, LinkedIn, Google+, YouTube, Yahoo, Vk, Instagram and others?***
> ***Is the data sufficiently protected during the transfer?***
> ***Is it encrypted and how?***

The Electronic Frontier Foundation has published a table depicting the current state of encryption practices concerning the major internet companies[238] which shows that DropBox, Facebook, Google, Microsoft, Sonic.net, Spideroak, Twitter and Yahoo do implement all encryption best practices suggested by the EFF. Other service providers like Amazon, Apple, FourSquare, LinkedIn, MySpace and tumblr, among others, do still lag behind when it comes to enabling encryption across their

[229] https://www.openssl.org/
[230] See Question 7
[231] http://msdn.microsoft.com/en-us/library/windows/desktop/aa380123%28v=vs.85%29.aspx
[232] http://support.microsoft.com/kb/977377
[233] https://technet.microsoft.com/library/security/MS14-066
[234] See Question 7
[235] http://www.unixwiz.net/techtips/iguide-ipsec.html
[236] http://tools.ietf.org/html/rfc6071
[237] http://www.kame.net/
[238] https://www.eff.org/deeplinks/2013/11/encrypt-web-report-whos-doing-what

networks. It must however be pointed out that this is a snapshot of the current state (Oct 2014) which is likely to change over time:

	Encrypts data center links	Supports HTTPS	HTTPS Strict (HSTS)	Forward Secrecy	STARTTLS
amazon	undetermined	limited	✗	undetermined	✗
Apple	undetermined	✓ (iCloud)	✗	undetermined	✗ (me.com, mac.com)
at&t	undetermined	undetermined	✗	undetermined	✗ (att.net)
comcast	undetermined	undetermined	✗	undetermined	✗ (comcast.net)
Dropbox	✓	✓	✓	✓	✓
facebook	✓ in progress	✓	✓ planned	✓	✓ (in progress, facebook.com)
foursquare	undetermined	✓	✓	undetermined	✗
Google	✓	✓	in progress for select domains, see notes	✓	✓
LinkedIn	✗ contemplating	✓	✓ planned 2014	✓ planned 2014	✓
Microsoft	✓ in progress	✓	✓ planned	✓ in progress	✓ (planned, outlook.com)
myspace	undetermined	✓	✗	undetermined	✗

	Encrypts data center links	Supports HTTPS	HTTPS Strict (HSTS)	Forward Secrecy	STARTTLS
Sonic.net	✓	✓	✓	✓ in progress	✓
SpiderOak	✓	✓	✓	✓ in progress	✓
@twitter	✓	✓	✓	✓	✓
tumblr.	✗	✓ planned Q2 2014	✓ planned 2014	✓	✗
verizon	undetermined	undetermined	✗	undetermined	✗ (verizon.net)
WordPress	undetermined	available	✗	undetermined	✗
YAHOO!	✓	✓ default for Mail; planned 2014 for all	✓ planned 2014	✓ yahoo.com; planned 2014 for all	✓ (yahoo.com)

Table 3: Encryption practices of major internet companies (source EFF)

2.8 Question 14

According to the Electronic Frontier Foundation (EFF), configuring browsers and website servers with the "Perfect forward secrecy" security option is a good solution to prevent access by mass surveillance organisations to encrypted content through cooperation with telecom operators (for traffic interception and collection purposes) and through service providers (to obtain the decryption keys).
Why are not all browsers and web-sites configured to support "Perfect forward secrecy"?
Why is it not extensively used today over the Internet?
What would be the advantages and disadvantages to deploy it massively?
Is this a good long-term solution?

Perfect Forward Secrecy (PFS) is very good solution[239] for preventing access for mass surveillance purposes. It uses new key generations for each session and if a security breach occurs, only the key being used in a particular session is compromised but never all the data transfer in previous sessions (as may happen when the server key is compromised in configurations that do not enable perfect forward secrecy) . Google supports PFS on their https sites.

The main technical problem with Perfect Forward Secrecy is its performance and throughput[240]. Compared to normal asymmetric algorithms it requires much more CPU cycles for being executed (15-27% throughput increase). This decrease in load-times of web pages would be the trade-off for achieving higher levels of security and is – obviously – the reason why PFS is not enabled on the majority of websites, with the exception of Gmail. An article in Computerworld lists some of those sites that have not enabled PFS, including a number of popular trade and retail sites, as well as large banks[241].

Netcraft, a UK based internet services company, has tested the availability of PFS for different browsers in June 2013 and found that "*whereas Google Chrome, Opera and Firefox were protected for approximately one third of connections, only a tiny fraction of Internet Explorer's SSL connections operated with PFS*" [242]. The full results of the study are documented in Figure 6.

[239] http://www.theregister.co.uk/2013/06/26/ssl_forward_secrecy/
[240] http://vincent.bernat.im/en/blog/2011-ssl-perfect-forward-secrecy.html
[241] http://www.computerworld.com/article/2473792/encryption/perfect-forward-secrecy-can-block-the-nsa-from-secure-web-pages--but-no-one-uses-it.html
[242] http://news.netcraft.com/archives/2013/06/25/ssl-intercepted-today-decrypted-tomorrow.html

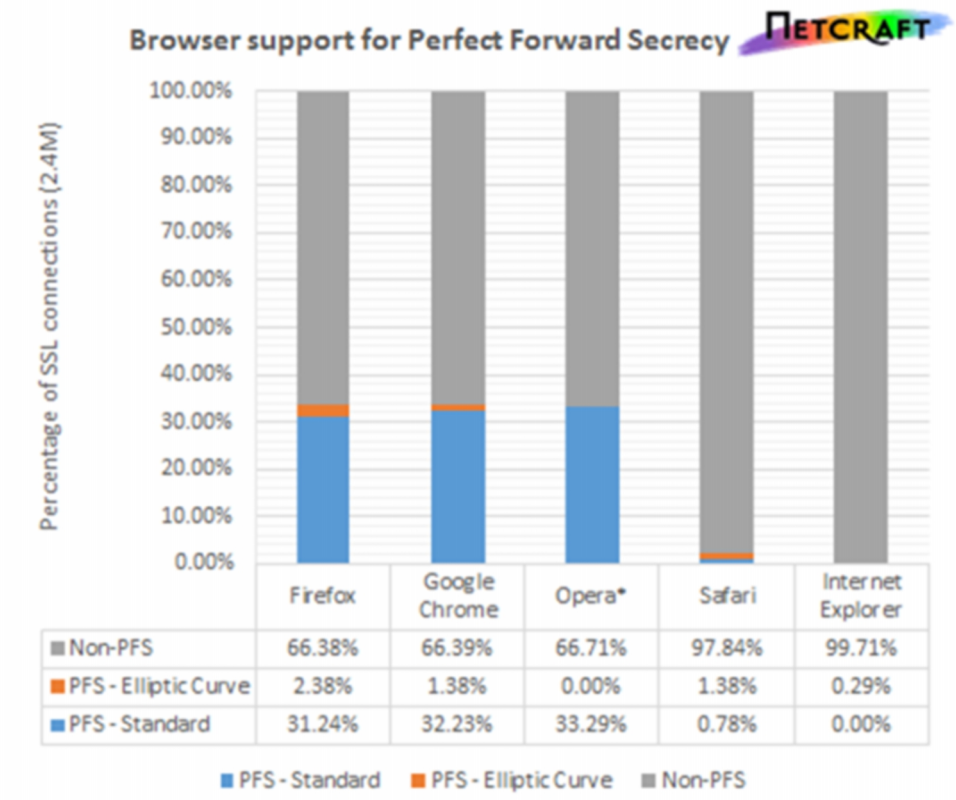


Figure 6: Browser support for PFS (as of June 2013 - source Netcraft)

The amount of information that is sent over the Internet is so huge that mass surveillance techniques usually collect all data and back it up, and only when it is necessary in retrospect the information is decrypted.

When a browser connects to a secure HTTPS server, it first sends some security parameters which establish security configuration of the connection. This configuration is called "cipher suite" and users generally don't have to worry about it because the browser configures the maximum security level by default. The Perfect Forward Secrecy (PFS) technique makes it more difficult to decrypt data in retrospect because each session has its own private key and no common private key like in SSL is used. The problem with SSL is that once an attacker gets hold of the common private key, all information exchanged in sessions that have been encrypted with this key can be decrypted.

PFS, in order to avoid retrospective decryption, separates authentication and private encryption keys. RSA can be used for authentication and the Diffie Helman[243] method for exchanging cryptographic keys. SSL provides two mechanisms to speed up connections[244]:

- **Session IDs:** Server and client remember their states through session id and communication is established using this parameter. The problem is that this session id is shared by front-end servers and usually saved on the client's hard disk.
- **Session tickets:** This is the default mode of OpenSSL and in this case the server sends the state to the client. Then, the client uses this state for authentication on other servers of the same cluster, avoiding full handshake (starting connection) and reducing connection time.

The principal reason why PFS is not used by default is its performance trade-off and, according to the comments of some experts, the lack of user demand prior to the revelations of Edward Snowden. However, the situation with respect to user demand for default encryption has dramatically changed

[243] http://en.wikipedia.org/wiki/Diffie%E2%80%93Hellman_key_exchange
[244] http://vincent.bernat.im/en/blog/2011-ssl-session-reuse-rfc5077.html

since the publication of these revelations and it can be expected that PFS, as one of the easiest solutions for preventing the breach of private SSL (server) keys, will be deployed in most servers in the future, likely making use of mechanisms that allow optimizing its performance.

2.9 Question 15

How long are "SSL/TLS and IPSEC decryption keys" archived by providers of Internet/Cloud/Social Network services in the US and the EU and what are their respective policies to make the decryption keys available to national security and/or law enforcement agencies?
Does the possession of these keys allow mass surveillance organisations to decrypt the traffic intercepted or taped from telecom providers?
What about the specific case of voice communication over terrestrial mobile telephony networks?
What about the specific case of data satellite transmissions?

Different reports in "ProPublica", the "New York Times" and the "Guardian" document NSA's and GCHQs practices and strategies for avoiding "*what they see as one of the biggest threats to their ability to access huge swathes of internet traffic – "the use of ubiquitous encryption across the internet"*."[245] These include "*covert measures*" to influence and gain control of international encryption standards, efforts to break encryption with brute force methods and "*collaboration with technology companies and internet service providers themselves*". The report lists the following facts that can be derived from the leaked documents:

- *A 10-year NSA program against encryption technologies made a breakthrough in 2010 which made "vast amounts" of data collected through internet cable taps newly "exploitable".*
- *The NSA spends $250m a year on a program which, among other goals, works with technology companies to "covertly influence" their product designs.*
- *The secrecy of their capabilities against encryption is closely guarded, with analysts warned: "Do not ask about or speculate on sources or methods."*
- *The NSA describes strong decryption programs as the "price of admission for the US to maintain unrestricted access to and use of cyberspace".*
- *A GCHQ team has been working to develop ways into encrypted traffic on the "big four" service providers, named as Hotmail, Google, Yahoo and Facebook.*

Other documents revealed by Edward Snowden show that the government has coerced companies to turn over information and keep official discretion on this practice (PRISM[246])

Mass surveillance on encrypted (also called "*secure*") web sites and other platforms are only possible if the respective surveillance organisation can acquire the private keys of the systems they want to spy on. Since these keys are stored in (text) files and files can be transferred and copied without accountability, it is relatively easy to leak and use private keys in different scenarios.

There are basically two ways of achieving this in real world scenarios:

1. Companies (either telecommunication or internet) are invited, paid or coerced (e.g. through court decisions) to give law enforcement and intelligence agencies access to (private) encryption keys or telephony metadata[247]

[245] http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security
[246] http://www.theguardian.com/us-news/prism
[247] http://www.forbes.com/sites/robertlenzner/2013/09/23/attverizonsprint-are-paid-cash-by-nsa-for-your-private-communications/

2. Surveillance organisations tap physical cables or illegally access telecom infrastructures beyond legal jurisdiction, as revealed by Snowden referring to programs labelled "Global Telecom Exploitation" and "Mastering the Internet".[248]

A foundation of solid cryptography is to you change encryption keys on a "regular" basis. However, the definition of what is "regular" depends on the administrators of the respective system. Public authentication keys are known to be maintained for long periods (up to decades), which in consequence enables retrospective decryption of encrypted communications – in case their metadata was saved – as soon as a private key counterpart is known. This weakness is addressed by the concept of PFS[249].

NIST has published a recommendation for key management [250] in which it states that the cryptoperiod of public signature-verification key "*may be on the order of several years, though due to the long exposure of protection mechanisms to hostile attack, the reliability of the signature is reduced with the passage of time*".

Table X documents the cryptoperiods recommended by NIST for different types of encryption keys:

Key Type	Cryptoperiod	
	Originator Usage Period (OUP)	Recipient Usage Period
1. Private Signature Key	1-3 years	
2. Public Signature Key	Several years (depends on key size)	
3. Symmetric Authentication Key	≤ 2 years	≤ OUP + 3 years
4. Private Authentication Key	1-2 years	
5. Public Authentication Key	1-2 years	
6. Symmetric Data Encryption Keys	≤ 2 years	≤ OUP + 3 years
7. Symmetric Key Wrapping Key	≤ 2 years	≤ OUP + 3 years
8. Symmetric and asymmetric RNG Keys	Upon reseeding	
9. Symmetric Master Key	About 1 year	
10. Private Key Transport Key	≤ 2 years	
11. Public Key Transport Key	1-2 years	
12. Symmetric Key Agreement Key	1-2 years	
13. Private Static Key Agreement Key	1-2 years	
14. Public Static Key Agreement Key	1-2 years	
15. Private Ephemeral Key Agreement Key	One key-agreement transaction	
16. Public Ephemeral Key Agreement Key	One key-agreement transaction	
17. Symmetric Authorization Key	≤ 2 years	
18. Private Authorization Key	≤ 2 years	
19. Public Authorization Key	≤ 2 years	

Table 4: Suggested cryptoperiods for key types (source NIST[250])

GSM network[251] encryption algorithms belong to A5 algorithms. Two adaptations of A5 were deployed and broken (A5/1 and A5/2.) and the current implementation (Kasumi) has demonstrated weaknesses[252],[253]. Therefore, current data transmission services such as 2G are vulnerable and the 3G

[248] http://siliconangle.com/blog/2013/07/19/how-the-nsa-taps-undersea-fiber-optic-cables/

[249] See also

Question 14

[250] http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57_part1_rev3_general.pdf
[251] http://www.sans.org/reading-room/whitepapers/telephone/gsm-standard-an-overview-security-317
[252] http://www.cs.technion.ac.il/users/wwwb/cgi-bin/tr-get.cgi/2006/CS/CS-2006-07.pdf
[253] http://www.ma.huji.ac.il/~nkeller/kasumi.ps

service has the major weakness of being susceptible to be lowered to 2G and then be attacked. The development of UMTS introduces an optional Universal Subscriber Identity Module (USIM), that customs a longer authentication key to provide greater security, as well as mutually authenticating the network and the user, whereas GSM only authenticates the user to the network (and not vice versa). The recommendation would therefore be to migrate to 4G 128 bits as soon as possible.

2.10 Question 16

What is the added value of the commercial solutions available to help users secure their on-line web, email, voice, chat communications over the Internet? This includes products and services such as Silent Circles, TOR, Ghostzilla, Disconnect, DoNotTrackme, Noscript, HTTPS everywhere, CryptoCat, Spider Oak, etc… To which extend are these commercial products effectively protecting the user or giving him a false sense of security?
Is the current generation of products mature?
What are the most promising commercial initiatives to date?
Are there some European solutions available?
What about European Initiatives such as LSEC and IPASCO and how mature are they compared to US ones?

Many commercial solutions for securing on-line communication are specific solutions for specific purposes. However, they do what they say and are gaining new adepts each day. The TOR network[254] is good example of how these solutions are spreading worldwide. Users require privacy and anonymity when surfing across Internet. The debate arises when this privacy is used to cover illegal or criminal activities, a situation that can be observed in what is called the "Deep Internet", a number of illegal websites that are only accessible though anonymizer services like, for instance the TOR network.

A main problem is the lack of awareness of these products by citizens. It seems that **most internet users are not really aware about how websites track their browsing behaviour when using conventional browsers**. **This is why above all a cultural and behavioural change is required when it comes to citizen's internet usage**, which not only includes creating awareness about possible security and privacy breaches, but also information dissemination on the tools and practices that can help to mitigate these risks.

For an extensive list of commercial solutions for securing communications over the internet please refer to Question 6.

2.11 Question 17

What is the reliability and quality of the encryption used over GSM voice networks in Europe, what is required to break it?
Are all voice communications encrypted by all operators?
Are there differences of reliability and quality of encryption between GSM network operators within Europe?
Any inter-operability issues?

There are different ways to attack GSM networks:

[254] https://www.torproject.org/

- Cryptographic attacks
- OTA attacks (Over The Air)
- Attacks from operators
- Memory corruption attacks
- Protocol vulnerabilities

The encryption of 3G networks covers only the connection between the end user phone and the operator base node (base station), the rest of communications are not encrypted.

GSM uses a variety of protocols: user identification, user authentication and data and voice transmission. Those algorithms are distributed among the GSM operator, the mobile phone and the SIM. Data and voice transmission encryption is implemented by A5 protocol, once the user is authenticated in the network. The A5 protocol has 3 different versions A5/1, A5/2 and A5/3. The A5/1 protocol used in GSM and responsible for maintaining the privacy of communications has been broken through reverse engineering. It is possible to decipher GSM communications in real time, intercepting voice and data from any user[255].

The 2G standard also has been broken and 3G connections can be handed over to 2G connections and thus be subject to attacks. This is, for instance, the strategy applied by 3G Cat, an IMSI/IMEI/TMSI Catcher[256] offered by the Israeli company Ability. As long as operators do not eliminate the 2G protocol, there will always be an access to eavesdrop voice and data during the transmission. Therefore, it is recommended to use the new encryption of 4G networks with a length of 128bits.

Other secure voice networks projects include the Secure Communication Interoperability Protocol (SCIP)[257], an application protocol designed by the US Department of Defense (DoD) and the NSA to operate independently of the characteristics of the underlying network channel. SCIP uses public cryptography (suite B) and is approved by NATO for protecting secret information.

2.12 Question 18

How about the quality and reliability of the raw encryption used over radio data networks such as 3G and 4G terrestrial networks and VSAT satellite data communications?

The protocol used for ciphering 3G is denominated Kasumi[258] (A5/3). Kasumi is based on an encryption technique, called Misty, that belongs to a subgroup of Feistel[259] techniques. KASUMI is an agile implementation of traditional Misty technique of 128 bits. The simplification of the algorithm has evoked new vulnerabilities, such as the so-called "sandwich attack", which can break an A5/3 encryption "*in less than two hours on a single PC*"[260]

4G is a new generation protocol which supports all Internet Protocol (IP) based communications. But the fact that it is based on the TCP/IP protocol implies that the 4G protocol inherits both, the potential and the weaknesses of TCP/IP, drawing particular attention to the large amount of hackers with IP knowledge.

[255] http://www.interceptors.com/decrypt-system/GSM.html
[256] http://www.interceptors.com/intercept-solutions/detects-parameters-3G-networks.html
[257] http://cryptome.org/2012/07/nsa-scip.pdf
[258] http://en.wikipedia.org/wiki/KASUMI
[259] http://en.wikipedia.org/wiki/Feistel_cipher
[260] http://eprint.iacr.org/2010/013.pdf

Neil Cook, chief technology officer at Cloudmark, a US based security company, claims that 4G is "*inherently less secure*" than previous mobile protocols[261]. "*Before 4G, all voice and data traffic between the user's device and the core of the network was encrypted and tightly-controlled by the mobile operator*". "*Many operators around the world, including some in Europe, have chosen to deploy 4G leaving the traffic between the core network and some or all of their cell sites un-encrypted*," said Patrick Donegan, senior analyst at Heavy Reading[261]. The portion of the network which is unencrypted by default is called backhaul. Most operators in Europe (not all) encrypt this network fragment with IPsec.

[261] http://www.telegraph.co.uk/technology/internet-security/10951812/4G-inherently-less-secure-than-3G.html

3 THEME 3 - EFFICIENCY OF COMMERCIAL PRODUCTS FOR DOING MASS SURVEILLANCE

3.1 Question 19

What are the tools commercially available for mass surveillance purposes to allow an organisation buying them to consolidate and analyse end-user meta-data intercepted?
What vendor provides what tools and what services to which organisations in which countries and for what purposes?
How legal is the commercialisation and subsequent use of such tools and services in Europe and elsewhere in the world?

Commercially available tools for surveillance purposes are classed by their vendors as advanced solutions for lawful communications data[262] (including both meta-data and content) interception, collection, processing and/or analysis for intelligence organizations, national security government agencies and Law Enforcement Agencies[263]. The vendors can provide this type of organizations with intelligence platforms and tools that collect, process and analyse both mass and targeted communications data[264],[265],[266].

As reported by the UN Special Rapporteur on the right to freedom of opinion and expression, "*the private sector has been complicit in developing technologies that enable mass or invasive surveillance in contravention of existing legal standards*"[267].

This report defines five different modalities of communications surveillance[267]:

1. Targeted communications surveillance
2. Mass communications surveillance
3. Access to communications data
4. Internet filtering and censorship
5. Restrictions on anonymity

Some examples of surveillance technology designed by the private vendors and utilized in Libya, Bahrain, the Syrian Arab Republic, Egypt and Tunisia are described in this report[268].

[262] Meta-data definition: information about an individual's communications (e-mails, phone calls and text messages sent and received, social networking messages and posts), identity, network accounts, addresses, websites visited, books and other materials read, watched or listened to, searches conducted, resources used, interactions (origins and destinations of communications, people interacted with, friends, family, acquaintances), and times and locations of an individual, including proximity to others. Ref: Human Rights Council, UN. A.HRC.23.4: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue, April 2013.

[263] http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA3-4842ENW.pdf

[264] http://www.nice.com/intelligence-lea

[265] http://blogs.wsj.com/digits/2011/11/21/surveillance-company-says-it-sent-fake-itunes-flash-updates-documents-show/

[266] http://www.verint.com/solutions/communications-cyber-intelligence/solutions/communications-interception/mass-interception/index

[267] Human Rights Council, UN. A.HRC.23.4: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue, April 2013.
http://www.ohchr.org/Documents/HRBodies/HRCouncil/RegularSession/Session23/A.HRC.23.40_EN.pdf

[268] European Parliament, Directorate-General for External Policies, Policy Department, After the Arab Spring: New Paths for Human Rights and the Internet in European Foreign Policy (2012), pp. 9-10.

Figure 7 provides an overview of surveillance deployment in Tunisia over the years; initially simply filtering websites and finally filtering and collecting all internet packets (using DPI based surveillance)[268].

State 1: 1997 – 2011: Web Filtering
Stage 2: 2003 – 2011: Email Filtering
Stage 3: 2007 – 2011: DPI based filtering & surveillance
Stage 4: 2010 – 2011: Hack Attacks

Figure 7: Timeline of surveillance deployment in Tunisia

Independently of censorship placed by Tunisian Government, communications surveillance was widely deployed both on the Internet and in mobile and fixed line telecommunications. Tunisie Telecom as well as private ISPs and mobile telecommunications operators were required to support Tunisian state in collecting citizens' communication data[268].

Mass surveillance technologies are often sold to countries in which there is a serious risk that they will be used to violate human rights, particularly those of human rights defenders, journalists or other vulnerable groups[268].

When talking about mass surveillance, this report says "*some States have the capability to track and record Internet and telephone communications on a national scale. By placing taps on the fibre-optic cables, through which the majority of digital communication information flows, and applying word, voice and speech recognition, States can achieve almost complete control of tele- and online communications. Such systems were reportedly adopted, for example, by the Egyptian and Libyan Governments in the lead-up to the Arab Spring*[269]."

Libya is another country where mass surveillance technologies were deployed and used during the Arab Spring. These surveillance systems were technologies provided by US as well as French companies[268]. In March 2012, the "Wall Street Journal" reported that the French technology company Amesys was selling a DPI based interception solution called Eagle to Libya[270]. The "Wall Street Journal" found an abandoned spy center in Tripoli where Eagle's training manuals and posters were discovered. One of the posters about the Eagle system said: "*Whereas many Internet interception systems carry out basic filtering on IP address and extract only those communications from the global flow* [lawful interception], *EAGLE Interception system analyses and stores all the communications from the monitored link* [massive interception]".

Another country that followed the Arab Spring Revolution was Syria. Although communications media had historically been restricted in this country, applying both censorship and surveillance, the Syrian Government started increasing the level of surveillance after the Revolution started[268]. The mass surveillance system was built with technology from European (Italy, France, and Germany) and US companies[271]: "*The project includes California-based NetApp Inc. (NTAP) storage hardware and software for archiving e-mails; probes to scan Syria's communications network from Paris-based Qosmos SA; and gear*

[269] Toby Mendel, Andrew Puddephatt, Ben Wagner, Dixi Hawtin, and Natalia Torres, Global Survey on Internet Privacy and Freedom of Expression, UNESCO Series on Internet Freedom (2012), p. 41.
[270] http://online.wsj.com/news/articles/SB10001424052970203961204577269391401776590?mg=reno64-wsj&url=http%3A%2F%2Fonline.wsj.com%2Farticle%2FSB10001424052970203961204577269391401776590.html
[271] http://www.bloomberg.com/news/2011-11-03/syria-crackdown-gets-italy-firm-s-aid-with-u-s-europe-spy-gear.html

from Germany's Utimaco Safeware AG (USA) that connects tapped telecom lines to Area's (Italy) monitoring-center computers".

According to Privacy International, in the ISS[272] (Intelligence Support Systems for lawful interception, electronic surveillance and cyber intelligence gathering) World Middle East 2014 conference, TeleStrategies[273] introduced the top ten Internet challenges faced by law enforcement, which are the following[274]:

- fibre bandwidth growth, which concerns DPI and fibre optical probes vendors.
- smartphones,
- encryption,
- social media monitoring,
- increasing data volume, being the solution based on Big Data analytics.
- tactical surveillance, related to geo-fenced calls and 3G and 4G antenna coverage.
- all IP (Long Term Evolution) infrastructure, replacing circuit switched networks.
- dark web [or deep web], which is enabled by TOR[275] and Bitcoin.
- dark email, being interesting two initiatives to an all encrypted email world: IETF's Prism-proof email protocols[276] and dark mail alliance.
- regulatory shortcomings, related to the laws that support the lawful intercept. According to the presentation, most of these laws are outdated and are voice circuit switched networks based.

According to TeleStrategies' presentation[277] the intelligent support systems for lawful interception, electronic surveillance and cyber intelligence gathering can be classified in eight product segments:

1. target/customer equipment linked to spyware, IT intrusion and forensics vendors
2. location and surveillance vendors
3. optical probe/DPI (Deep Packet Inspection) vendors
4. mediation/DPI vendors
5. data retention vendors
6. monitoring centre vendors
7. secure communications vendors
8. intelligence centre analytics such as Big Data.

[272] http://www.issworldtraining.com/
[273] http://www.telestrategies.com/
[274] https://www.documentcloud.org/documents/1215458-1299-telestrategies-presentation-challenges.html#document/p46/a178126
[275] https://www.torproject.org/
[276] http://prismproof.org/index.html
[277] https://www.documentcloud.org/documents/1215458-1299-telestrategies-presentation-challenges.html#document/p46/a178126

Figure 8 depicts this product categorization detailing the main point of surveillance over the network architecture.

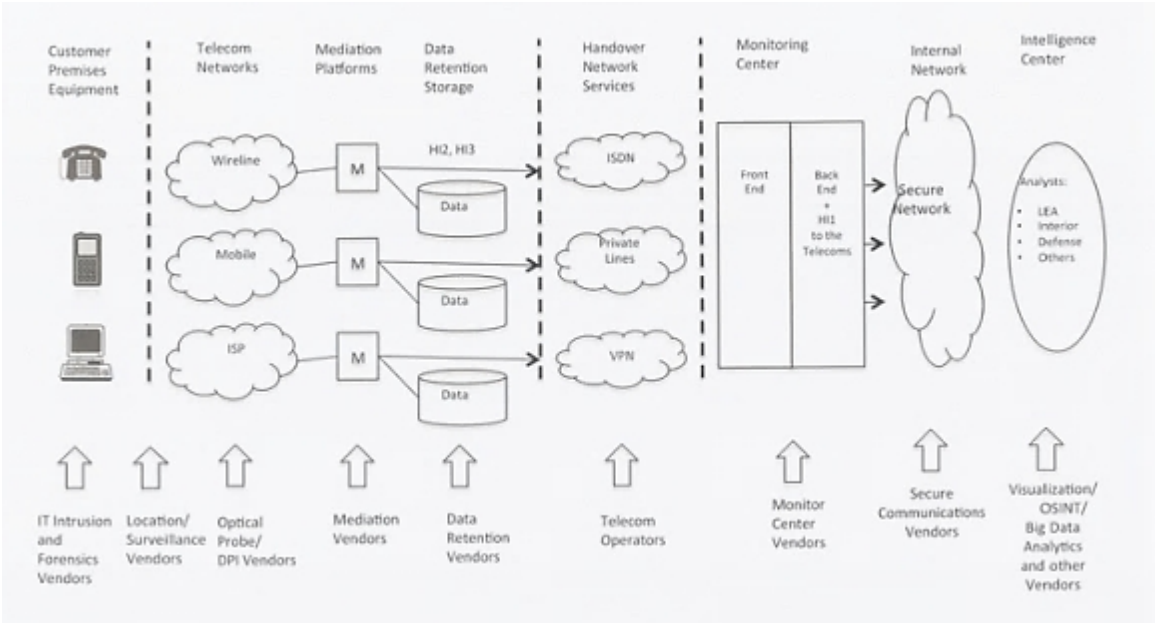


Figure 8: Lawful Interception, Intelligence Gathering, Telecom Networks and ISS Vendors (source Telestrategies)

Figure 9 shows the most important vendors of intelligent support systems for lawful interception and electronic surveillance classified by vendors groups[278]: IT intrusion vendors, location and surveillance vendors, optical probe and DPI vendors, mediation data retention and monitoring centre vendors, and intelligence centre vendors.

[278] https://www.documentcloud.org/documents/1215458-1299-telestrategies-presentation-challenges.html#document/p46/a178126

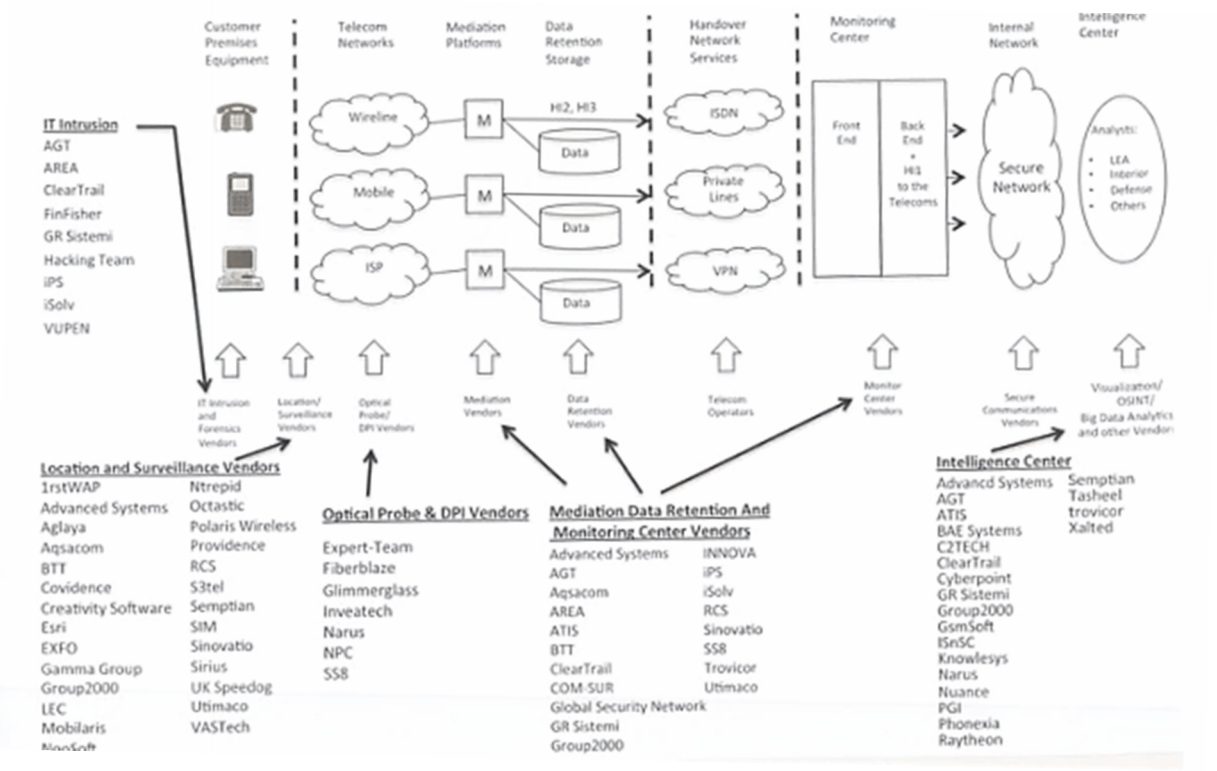


Figure 9: Lawful Interception and Intelligence Gathering Vendors (source Telestrategies)

The surveillance vendors highlighted below have been accused of providing countries that violate human rights with surveillance technology[279,280].

The author of this report[281] claimed that "*while the majority of these companies* [surveillance vendors] *claim to sell their products to a restricted client base of law enforcement, military, and intelligence agencies, this report shows another example of commercial network intrusion tools being used against dissidents in countries with poor human rights records*."

Gamma Group is an UK/Germany based[280] company that provides National and State Intelligence Departments and Law Enforcement Agencies with advanced technical surveillance and monitoring solutions and advanced government training as well as international consultancy[282].

FinFisher is a commercial intrusion kit distributed by Gamma Group and is regarded as the most advanced surveillance solution in the market. This product includes trojans to infect PCs, mobile phones, other consumer electronics and servers, as well as technical consulting. The FinSpy spyware software can be considered as the FinFisher's remote monitoring product. FinSpy captures information from an infected computer, such as passwords and Skype calls, and sends the information to a FinSpy command & control server[281].

In 2013, the Citizen lab in the University of the Toronto published a map of Finfisher proliferation around the world, which is documented in Figure 10. This map shows the results of scanning for characteristics of FinFisher Command and Control servers[283]. FinFisher servers have been found in 25

[279] http://surveillance.rsf.org/en/

[280] http://surveillance.rsf.org/en/wp-content/uploads/sites/2/2013/03/enemies-of-the-internet_2013.pdf

[281] Morgan Marquis-Boire, Bill Marczak, Claudio Guarnieri and John Scott-Railton. For Their Eyes Only: The Commercialization of Digital Spying. Citizen Lab and Canada Centre for Global Security studies, University of Toronto. May 2013. https://citizenlab.org/storage/finfisher/final/fortheireyesonly.pdf

[282] https://www.gammagroup.com/Default.aspx

[283] https://citizenlab.org/storage/finfisher/maps/5-clicktwice-highrez.jpg

countries: Australia, Bahrain, Bangladesh, Brunei, Canada, Czech Republic, Estonia, Ethiopia, Germany, India, Indonesia, Japan, Latvia, Malaysia, Mexico, Mongolia, Netherlands, Qatar, Serbia, Singapore, Turkmenistan, United Arab Emirates, United Kingdom, United States and Vietnam.

However, the authors of the investigation pointed out that list of servers might be incomplete due to the large diversity of ports used by FinFisher servers. Furthermore, they said that discovery of a FinSpy command and control server in a given country is not a sufficient indicator to conclude the use of FinFisher by that country's law enforcement or intelligence agencies. In some cases, servers were found running on facilities provided by commercial hosting providers that could have been purchased by actors from any country[281].

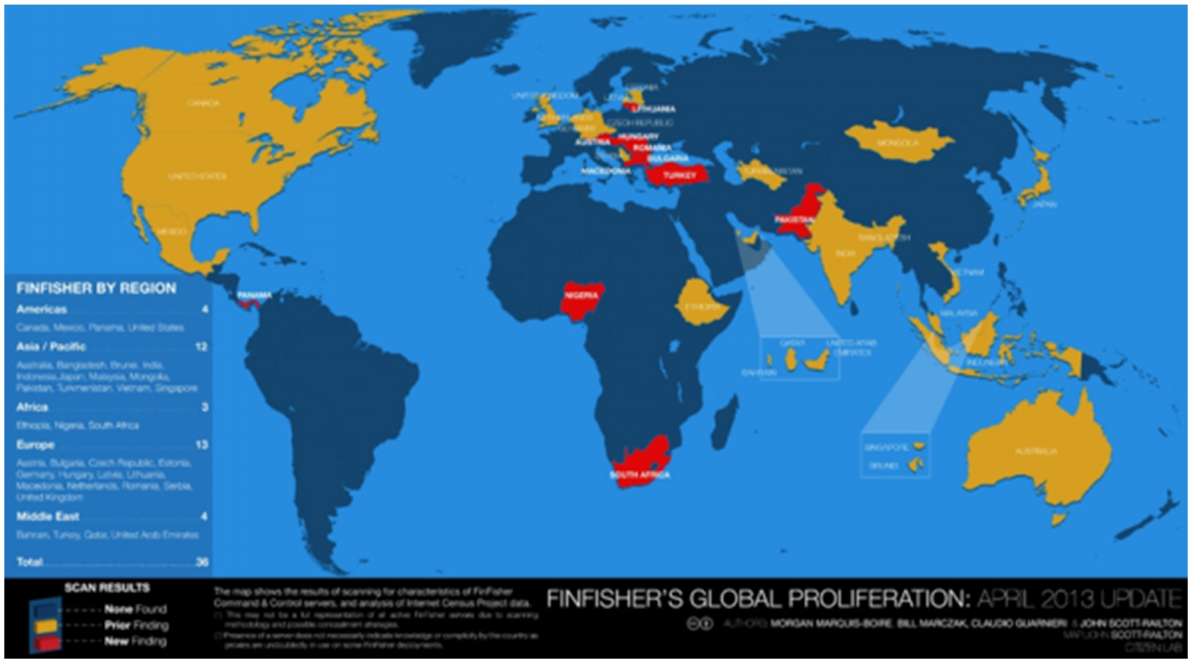


Figure 10: Global Proliferation of FinFisher surveillance suite

Hacking Team[284] is an Italian based surveillance oriented company that offers "offensive" capabilities for law enforcement agencies and intelligence agencies. They state about their Remote Control System solution that[285]: "*In modern digital communications, encryption is widely employed to protect users from eavesdropping. Unfortunately, encryption also prevents law enforcement and intelligence agencies from being able to monitor and prevent crimes and threats to the country security. Remote Control System (RCS) is a solution designed to evade encryption by means of an agent directly installed on the device to monitor. Evidence collection on monitored devices is stealth and transmission of collected data from the device to the RCS server is encrypted and untraceable.*"

Their customer policy claims that[286] "*we don't sell products to governments or to countries blacklisted by the U.S., E.U., U.N., NATO or ASEAN. We review potential customers before a sale to determine whether or not there is objective evidence or credible concerns that Hacking Team technology provided to the customer will be used to facilitate human rights violations.*"

However several media reports and IT security experts have found traces of their software in countries with a questionable position about human rights such as Morocco and United Arab Emirates[280].

[284] http://www.hackingteam.it/
[285] http://www.hackingteam.it/index.php/remote-control-system
[286] http://www.hackingteam.it/index.php/customer-policy

Blue Coat[287] is a US based company that is specialized in online security but it is well known in the surveillance market for its DPI technology based equipment[280]. This surveillance solution was supposed to be sold to countries such as Syria and Burma (Myanmar) [280].

IT researchers from the Citizen Lab of the University of Toronto carried out an investigation on Blue Coat devices and they highlighted that Blue Coat solutions, capable of filtering, censorship and surveillance are being used around the world[288]. They found 61 Blue Coat ProxySG devices and 316 Blue Coat PacketShaper appliances in the following locations[288]:

- Blue Coat ProxySG: Egypt, Kuwait, Qatar, Saudi Arabia, the UAE.
- PacketShaper: Afghanistan, Bahrain, China, India, Indonesia, Iraq, Kenya, Kuwait, Lebanon, Malaysia, Nigeria, Qatar, Russia, Saudi Arabia, South Korea, Singapore, Thailand, Turkey, and Venezuela.

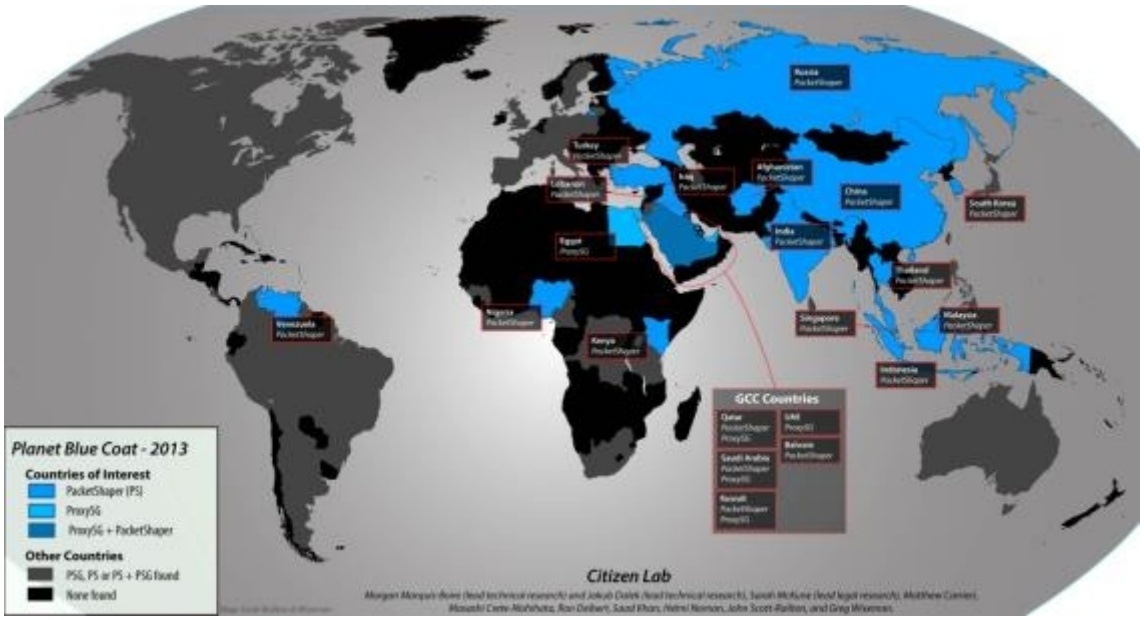


Figure 11: Blue Coat gobal deployment

Table 5 shows a list of catalogues that includes a great variety of mass surveillance vendors as well as information related to their products and in which countries these products have been used.

Catalogue	# Surveillance vendors	# Products	#Countries	#Documents
The Spyfiles[289]	95	78	-	287
Surveillance Industry Index[290]	375	11 (technologies)	66	1305
Bugged Planet [291]	136[292]			

[287] https://www.bluecoat.com/
[288] https://citizenlab.org/2013/01/planet-blue-coat-mapping-global-censorship-and-surveillance-tools/
[289] https://wikileaks.org/the-spyfiles.html
[290] https://www.privacyinternational.org/sii
[291] http://buggedplanet.info/index.php?title=Main_Page

Reporters without borders[293][294] [280]	5		16	

Table 5: Catalogues of surveillance vendors and products

The Spyfiles catalogue includes an interesting interactive map[295] which shows the use of different surveillance technologies in a number of countries around the world. Figure 12 shows the respective record of surveillance companies and technologies in the UK.

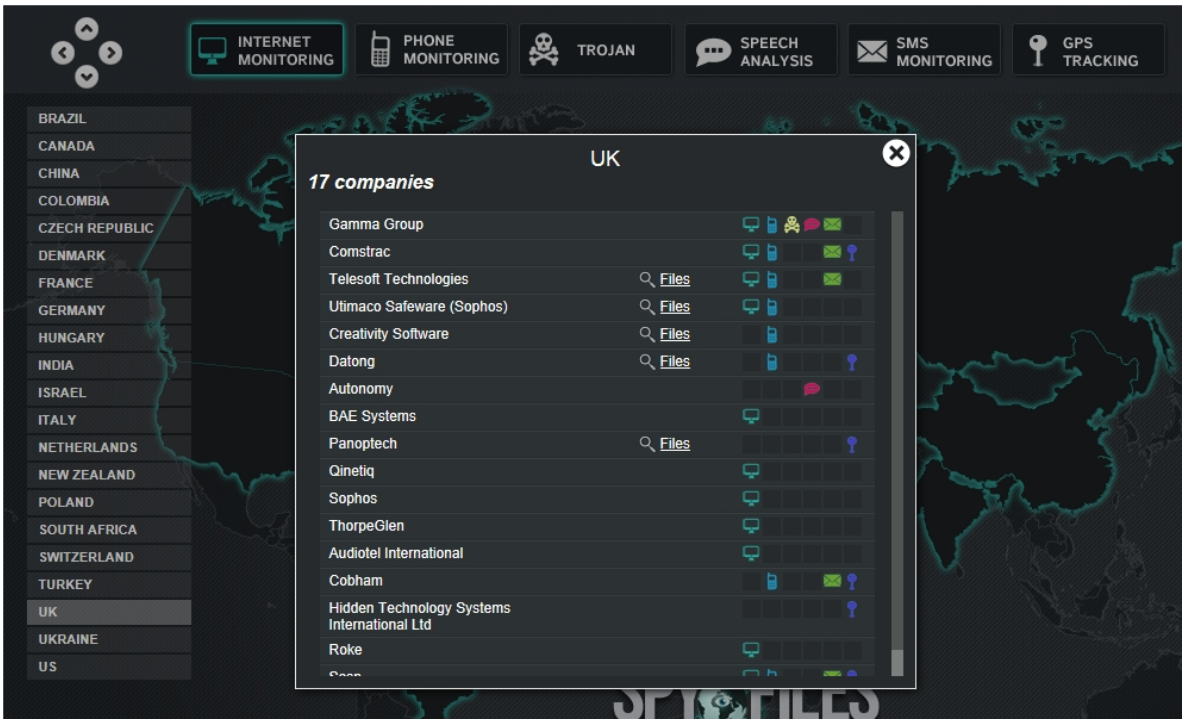


Figure 12: Surveillance companies and technologies in the UK (source Wikileaks Spyfiles-map)

Legal Aspects

In December 2012, the European Union banned the export of information technology that can be used by repressive regimes to censor information, conduct mass surveillance and track people's movements[296]. The USA has also established sanctions against a list of countries, using the blocking of assets and trade restrictions to accomplish foreign policy and national security goals[297]. The list of countries includes Iran, North Korea, Sudan and Syria among others.

A total of 42 countries[298] already negotiated the inclusion of surveillance technology into the most comprehensive international treaty on export controls[280], the Wassenaar Arrangement (WA)[299]. In 2013 new export controls in surveillance and law enforcement/intelligence gathering tools and IP network

[292] Each vendor information includes information about products, articles, events and situations released on the media or in the Internet.
[293] http://surveillance.rsf.org/en
[294] http://en.rsf.org/countries-under-surveillance,32.html
[295] https://wikileaks.org/The-Spyfiles-The-Map.html
[296] http://www.europarl.europa.eu/news/en/news-room/content/20121207IPR04411/html/Ban-exports-of-information-technology-to-repressive-regimes
[297] http://www.treasury.gov/resource-center/sanctions/Programs/Pages/Programs.aspx
[298] http://www.wassenaar.org/participants/index.html
[299] http://www.wassenaar.org/introduction/origins.html

surveillance systems or equipment were agreed among the WA participating States[300]. As stated in the "Guidelines and Procedures including the Initial elements"[301] of the WA*: "The Wassenaar Arrangement has been established in order to contribute to regional and international security and stability, by promoting transparency and greater responsibility in transfers of conventional arms and dual-use goods and technologies, thus preventing destabilising accumulations. Participating States will seek, through their national policies, to ensure that transfers of these items*[302] *do not contribute to the development or enhancement of military capabilities which undermine these goals, and are not diverted to support such capabilities."*

The British Government indeed expressed its commitment to putting in place export controls on telecommunication surveillance equipment "*for which there is a reasonable expectation that it might be used to restrict freedom of expression on the internet*"[303] through the WA mechanism.

However, several companies from the surveillance industry located their centre to export their surveillance technologies around the world in Switzerland, which is a participating State in the WA[304]. But in February 2014 most of them cancelled the export applications of surveillance technology, including all applications for the export of trojans and technologies for internet monitoring, due to a Swiss government inquiry into the human rights impact of the commercial surveillance trade[305].

In October 2014, the European Commission added specific forms of surveillance technology to the EU control lists on dual use items[306]. This control enables European Commission to identify the accountability of selling, transferring and exporting surveillance technology by the private sector.

Surveillance is subject to different international and national legal norms:

General public international law

As stated by Emeritus Professor Douwe Korff:

> *Surveillance by one state over the Internet activities and electronic communications of citizens and officials of another state with which the first state is not at war at that time, without the express consent of the other state, and which involve illegal activities by agents of the first state perpetrated within the territory of the other state, is a violation of the sovereignty of the targeted state. This is a rule of primary international law*[307].

Moreover,

> *…surveillance of citizens and officials of one state-party to an international human rights treaty by agents of another state-party to that treaty, from the territory of the latter state, but which violates the obligations of the latter state party under that treaty, not only violates that treaty but (since it harms the interests of the targeted state and its officials and citizens) also constitutes an internationally unlawful act against the state whose citizens and officials are affected. That is a rule of secondary international law.*

[300] http://www.wassenaar.org/publicdocuments/2013/WA%20Plenary%20Public%20Statement%202013.pdf
[301] http://www.wassenaar.org/guidelines/docs/Guidelines%20and%20procedures%20including%20the%20Initial%20Elements.pdf
[302] http://www.wassenaar.org/controllists/index.html
[303] https://www.privacyinternational.org/news/press-releases/british-government-welcomes-foreign-affairs-committee-recommendation-to-control
[304] https://www.privacyinternational.org/news/blog/after-gamma-revelations-switzerland-begins-to-debate-export-of-surveillance-tech
[305] https://www.privacyinternational.org/news/blog/surveillance-companies-ditch-switzerland-but-further-action-needed
[306] https://www.privacyinternational.org/news/press-releases/eu-catches-up-takes-steps-to-control-export-of-intrusion-spyware-ip-monitoring
[307] Expert Opinion prepared for the Committee of Inquiry of the Bundestag into the "5EYES" global surveillance systems revealed by Edward Snowden, Committee Hearing, Paul-Löbe-Haus, Berlin, 5 June 2014.

International and European human rights law

International human rights law is mainly treaty-based. The International Covenant on Civil and Political Rights (ICCPR) is the International human rights treaty and the European Convention on Human Rights (ECHR) is the European human rights treaty.

> *The provisions in the ECHR and the ICCPR (on the rights if private life, freedom of expression, freedom of information, and freedom of association) all stipulate or imply that those rights can only be restricted or interfered with on the basis of "law"; and that such restrictions or interferences must serve a "legitimate aim", and must be "necessary" to achieve that aim.*

These aims include public safety, prevention of crime, protection of morals and of the rights of others, and national security. The UN has published a declaration on the use of surveillance technologies and their impact on human rights, specifically on the right of privacy[308]:

> *In most States**, legal standards are either non-existent or inadequate to deal with the modern communications surveillance environment**. As a result, States are increasingly seeking to justify the use of new technologies within the ambits of old legal frameworks, without recognizing that the expanded capabilities they now possess go far beyond what such frameworks envisaged. In many countries, this means that vague and broadly conceived legal provisions are being invoked to legitimize and sanction the use of seriously intrusive techniques[267].*
>
> *When conducted in compliance with the law, including international human rights law, surveillance of electronic communications data can be a necessary and effective measure for legitimate law enforcement or intelligence purposes. There may be legitimate reasons for a State to require that an information and communications technology company provide user data; however, when a company supplies data or user information to a State in response to a request that contravenes the right to privacy under international law, **a company provides mass surveillance technology or equipment to States without adequate safeguards in place or where the information is otherwise used in violation of human rights, that company risks being complicit in or otherwise involved with human rights abuses**.*
>
> *Practices in many States have, however, revealed a **lack of adequate national legislation and/or enforcement, weak procedural safeguards, and ineffective oversight,** all of which have contributed to a **lack of accountability** for arbitrary or unlawful interference in the right to privacy.*
>
> *Mass interception technology eradicates any considerations of proportionality, enabling indiscriminate surveillance. It enables the State to copy and monitor every single act of communication in a particular country or area, without gaining authorization for each individual case of interception.*
>
> *The very existence of a mass surveillance programme thus creates an **interference with privacy**.*
>
> *In response to the increased data flows across borders and the fact the majority of communications are stored with foreign third party service providers, a number of States have begun to adopt laws that purport to authorize them to conduct extra-territorial surveillance or to intercept communications in foreign jurisdictions. This raises serious concern with regard to the extra-territorial commission of human rights violations and the inability of individuals to know that they might be subject to foreign surveillance, challenge decisions with respect to foreign surveillance, or seek remedies.*
>
> ***Inadequate legal standards increase the risk of individuals being exposed to violation of their human rights**, including the right to privacy[309] and the right to freedom of expression.*

[308] A/HRC/27/37: The right to privacy in the digital age. Report of the Office of the United Nations High Commissioner for Human Rights. 30 June 2014.

[309] At both the international and regional levels, privacy is also unequivocally recognized as a fundamental human right. The right to privacy is enshrined by the Universal Declaration of Human Rights (art. 12), the International Covenant on Civil and Political Rights (ICCPR, art. 17), the Convention on the Rights of the Child (art. 16), and the International Convention on the Protection of All Migrant Workers and Members of Their

States must take measures to prevent the commercialization of surveillance technologies, paying particular attention to research, development, trade, export and use of these technologies considering their ability to facilitate systematic human rights violations.

The UN states that when surveillance is conducted under the law, the data collected can be necessary and effective for the law enforcement purposes. But mass interception technology eliminates any considerations of proportionality, enabling the State to collect personal data without gaining authorization for each individual case of interception and this situation creates interference with privacy. Moreover, a company that provides mass surveillance technology to States without adequate safeguards in place or where the information is otherwise used in violation of human rights, the company risks being complicit in or otherwise involved with human rights violation[310].

Generally, a judicial authorization is required for performing communications surveillance in European countries. But as described by the UN, in many cases national intelligence agencies also enjoy blanket exceptions to the requirement for judicial authorization. *For example, German law allows warrantless automated wiretaps of domestic and international communications by the State's intelligence services for the purposes of protecting the free democratic order, existence or security of the State. In Sweden, the Law on Signals Intelligence in Defense Operations authorizes the Swedish intelligence agency to intercept without any warrant or court order all telephone and Internet traffic that takes place within Sweden's borders*[267].

And *in the United States, the Foreign Intelligence Surveillance Act empowers the National Security Agency to intercept communications without judicial authorization where one party to the communication is located outside the United States, and one participant is reasonably believed to a member of a State-designated terrorist organization*[267].

The UN declares that *mass surveillance technology can contribute to the suppression and prosecution of acts of terrorism does not provide an adequate human rights law justification for its use. The fact that something is technically feasible, and that it may sometimes yield useful intelligence, does not mean that it is either reasonable or lawful (in terms of international or domestic law)*[311].

The Emeritus Professor Douwe Korff summarized the European Court of Human Rights considerations and minimum requirements relating to State surveillance[312]:

ECtHR CONSIDERATIONS & MINIMUM REQUIREMENTS RELATING TO SURVEILLANCE:

The case-law of the ECtHR shows the following **considerations and requirements of European human rights law relating to surveillance**:

- A system of secret surveillance for the protection of national security may undermine or even destroy democracy under the cloak of defending it.
- The mere existence of legislation which allows a system for the secret monitoring of communications entails a threat of surveillance for all those to whom the legislation may be applied.
- In view of these risks, there must be adequate and effective guarantees against abuse.
- The first of these is that such systems must be set out in statute law, rather than in subsidiary rules, orders or manuals. The rules must moreover be in a form which is open to public scrutiny and knowledge. Secret, unpublished rules in this context are fundamentally

Families (art. 14). At the regional level, the right to privacy is protected by the European Convention on Human Rights (art. 8) and the American Convention on Human Rights (art. 11)

[310] https://www.privacyinternational.org/news/blog/un-privacy-report-a-game-changer-in-fighting-unlawful-surveillance

[311] UN. A/69/397: Promotion and protection of human rights and fundamental freedoms while countering terrorism. 23 September 2014.

[312] http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/dv/note_korff_/note_korff_en.pdf

contrary to the Rule of Law; surveillance on such a basis would ipso facto violate the Convention.

The following are the **"minimum safeguards" that should be enshrined in such (published) statute**:

- the offences and activities in relation to which surveillance may be ordered should be spelled out in a clear and precise manner;
- the law should clearly indicate which categories of people may be subjected to surveillance;
- there must be strict limits on the duration of any ordered surveillance;
- there must be strict procedures to be followed for ordering the examination, use and storage of the data obtained through surveillance;
- there must be strong safeguards against abuse of surveillance powers, including strict purpose/use-limitations (e.g., preventing the too-easy disclosure of intelligence data for criminal law purposes) and strict limitations and rules on when data can be disclosed by NSAs to LEAs, etc.;
- there must be strict rules on the destruction/erasure of surveillance data to prevent surveillance from remaining hidden after the fact;
- persons who have been subjected to surveillance should be informed of this as soon as this is possible without endangering national security or criminal investigations, so that they can exercise their right to an effective remedy at least ex post facto; and
- the bodies charged with supervising the use of surveillance powers should be independent and responsible to, and be appointed by, Parliament rather than the Executive.

Under the ECHR, these principles must be applied to anyone who is affected by surveillance measures taken by any Council of Europe Member State.

In addition, European States have a **"positive obligation" to protect their citizens from surveillance contrary to the above, perpetrated by any other State. A fortiori, they are under a legal obligation not to actively support, participate or collude in such surveillance**.

International and European data protection law

In 1995, the first (and still the main) EC directive on data protection[313] was adopted, followed by a specialised, subsidiary directive on privacy and electronic communications in 2002 (the "e-Privacy Directive"[314]), and, in 2006, by the so-called "Data Retention Directive"[315], which is technically an amendment to the e-Privacy Directive.

In Europe, data protection is seen as an essential pre-requisite for the protection of other freedoms, including freedom of thought and freedom of expression. But the law allows for exceptions which must be:

[313] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 23 November 1995, OJ L.281, p. 31ff.

[314] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications, commonly referred to as the e-Privacy Directive), 31 July 2002, OJ L 201, p. 37ff.

[315] Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, 13 April 2006, OJ L105, p. 54ff.

a necessary, appropriate and proportionate measure within a democratic society to safeguard national security (i.e. State security), defence, public security, and the prevention, investigation, detection and prosecution of criminal offences or of unauthorised use of the electronic communication system[312].

But in a crucial recent judgment[316], *the Court of Justice ruled that the Data retention Directive failed to meet these requirements, because it imposed departures from the core data protection principles that were not proportionate to the stated aim of the Directive*[307]. *The Directive was fundamentally flawed because it both lacked sufficiently "clear and precise rules" to circumscribe the capturing of data, and "sufficient safeguards against abuse", the Court concluded that in adopting the Data Retention Directive in the form it did, "the EU legislature has exceeded the limits imposed by compliance with the principle of proportionality".*

3.2 Question 20

What type of data input do commercial mass surveillance tools require in practice?
How easy and user-friendly are these tools to use to produce the desired results?
What are concrete examples of the type reports or queries that can be issued using the most powerful tools available on the market?
What is the level of granularity of the analysis that can be achieved practically?
What is the level of expertise and skills needed by the operators using such commercial tools?

Thanks to the publication of internal documents of the Gamma Group[317], vendor of the FinFisher surveillance suite, by a hacker who was able to retrieve these documents from their internal network, it is possible to answer this question without having direct access to the mentioned software application. **Due to the lack of comparable operational information for other surveillance software documented in Question 19, much of the information concerning Question 20 is constricted to FinFisher**, although it can be expected that similar tools do also provide similar capabilities and feature sets.

The FinFisher product portfolio presentation[318] lists various components along the three product vectors "Remote Monitoring & Infection Solutions", "Tactical IT Intrusion Portfolio" and "IT Intrusion Training Program" (see Figure 13).

[316] Judgment in Joined Cases C-293/12 and C-594/12, Digital Rights Ireland and Seitlinger and Others, of 8 April 2014, available at:
http://curia.europa.eu/juris/document/document.jsf?text=&docid=150642&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=12322
For an early analysis, see:
http://eulawanalysis.blogspot.co.uk/2014/04/the-data-retention-judgment-cjeu.html
[317] https://wikileaks.org/spyfiles4/index.html
[318] https://wikileaks.org/spyfiles/files/0/299_GAMMA-201110-FinFisher_Product_Portfolio-en.pdf

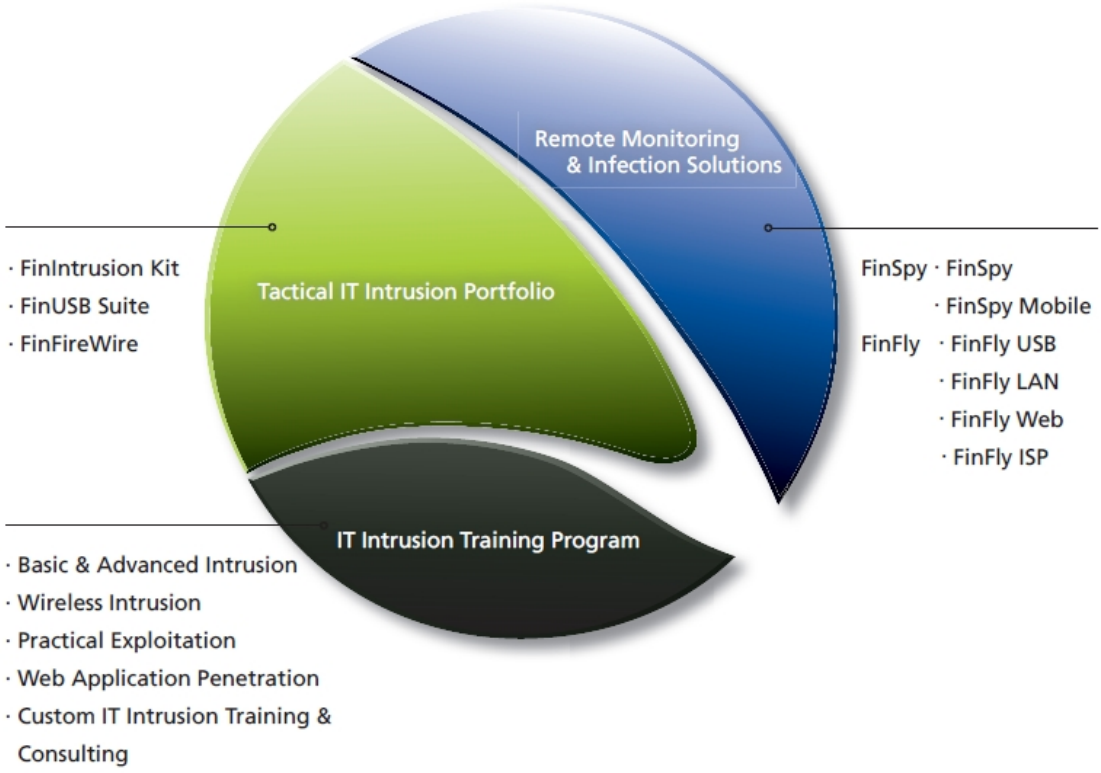


Figure 13: Components of FinFisher product portfolio (source Gamma Group)

The features of the different components are documented in the following Table 6:

FinIntrusion Kit *The FinIntrusion Kit is an **up-to-date and covert** operational Kit that can be used for most common **IT Intrusion Operations** in defensive and offensive areas*	• Discovers **Wireless LANs (802.11) and Bluetooth® devices** • Recovers WEP (64 and 128 bit) Passphrases **within 2-5 minutes** • **Breaks WPA1 and WPA2** Passphrases using Dictionary Attacks • Actively monitors Local Area Network (Wired and Wireless) and **extracts Usernames and Passwords even for TLS/SSL-encrypted sessions** • Emulates **Rogue Wireless Access-Point** (802.11) • Remotely **breaks into Email Accounts** using Network-, System- and Password-based Intrusion Techniques • **Network Security Assessment** and Validation

FinUSB Suite
 	
*The FinUSB Suite is a flexible product that enables Law Enforcement and Intelligence Agencies to quickly and securely extract forensic information from computer systems **without the requirement of IT-trained Agents**.*
	• Optimized for **Covert Operations**
	• Easy usability through **Automated Execution**
	• **Secure Encryption** with RSA and AES
	• Extraction of **Usernames and Passwords** for all common software like:
	  ○ Email Clients
	  ○ Messengers
	  ○ Browsers
	  ○ Remote Administration Tools
	• **Silent Copying of Files** (Search Disks, Recycle-Bin, Last opened/edited/created)
	• Extracting **Network Information** (Chat Logs, Browsing History, WEP/WPA(2) Keys, …)
	• Compilation of **System Information** (Running/Installed Software, Hard-Disk Information, …)

FinFireWire	• **Unlocks User-Logon** for every User-Account

*FinFireWire enables the Operator to quickly and covertly **bypass the password-protected** screen and access the Target System without leaving a trace or harming essential forensic evidence.*	• Unlocks **Password-Protected Screensaver** • Full Access to **all Network Shares** of User • **Dumps full RAM** for Forensic analysis • Enables live forensics **without rebooting** the Target System • User password is **not changed** • Supports **Windows, Mac and Linux systems** • Works with **FireWire/1394, PCMCIA and Express Card**
FinSpy *FinSpy is a field-proven Remote Monitoring Solution that enables Governments to face the current challenges of **monitoring Mobile and Security-Aware Targets** that regularly **change location**, use **encrypted and anonymous communication** channels and **reside in foreign countries**.*	• Bypassing of 40 regularly tested Antivirus Systems • **Covert Communication** with Headquarters • Full **Skype Monitoring** (Calls, Chats, File Transfers, Video, Contact List) • Recording of **common communication** like Email, Chats and Voice-over-IP • **Live Surveillance** through Webcam and Microphone • **Country Tracing** of Target • **Silent extracting of Files** from Hard-Disk • **Process-based Key-logger** for faster analysis • **Live Remote Forensics** on Target System • **Advanced Filters** to record only important information • Supports most common Operating Systems (**Windows, Mac OSX** and **Linux**)
FinSpy Mobile *FinSpy Mobile is closing the gap of interception capabilities for Governments for most common **smart phone platforms.***	• **Covert Communications** with Headquarters • Recording of **common communications** like Voice Calls, SMS/MMS and Emails • **Live Surveillance** through silent Calls • **File Download** (Contacts, Calendar, Pictures, Files) • **Country Tracing** of Target (GPS and Cell ID) • Full Recording of all **BlackBerry Messenger communications** • Supports most common Operating Systems: **Windows Mobile, iOS (iPhone), BlackBerry and Android**
FinFly USB *Once the FinFly USB is inserted into a computer, it **automatically installs the configured software** with little or no user-interaction and **does not require IT-trained Agents** when being used in operations. The FinFly USB can be used against **multiple systems** before being returned to Headquarters.*	• **Covertly installs Remote Monitoring Solution** on insertion in Target System • **Little or no user-interaction** is required • Functionality can be **concealed by placing regular files** like music, video and office documents on the device • Infection of **switched off Target System** when **booting from USB** • Hardware is a **common and non-suspicious USB device**
FinFly LAN *FinFly LAN was developed to deploy a Remote Monitoring Solution covertly on Target Systems in Local Area Networks (Wired and Wireless/802.11). It is able*	• **Discovers all Computer Systems** connected to Local Area Network • Works in **Wired and Wireless (802.11)** Networks • Can be combined with FinIntrusion Kit for **covert Network Access** • Hides Remote Monitoring Solution in **Downloads of Targets** • Injects Remote Monitoring Solution as **Software Updates** • Remotely installs Remote Monitoring Solution through **Websites visited by the Target**

*to **infect Files that are downloaded** by the Target on-the-fly, infect the Target by **sending fake Software Updates** for popular Software or infect the Target by **injecting the Payload into visited Websites**.*	
FinFly Web *FinFly Web is designed to provide **remote and covert** infection of a Target System by using a wide range of **web-based attacks**. FinFly Web provides a **point-and-click interface**, enabling the Agent to easily **create a custom infection code** according to selected modules.*	• **Fully-Customizable** Web Modules • Can be covertly **installed into every Website** • Full integration with **FinFly LAN** and **FinFly ISP** to deploy even inside popular Websites like Webmail, Video Portals and more • Installs Remote Monitoring Solution **even if only email address is known** • Possibility to target every person visiting **configured Websites**
FinFly ISP *FinFly ISP is a strategic, **countrywide, as well as a tactical** (mobile) solution that can be **integrated into an ISP's Access and/or Core Network** to remotely install the Remote Monitoring Solution on selected Target Systems.*	• Can be installed inside the **Internet Service Provider Network** • Handles **all common Protocols** • Selected Targets by **IP address or Radius Logon Name** • Hides Remote Monitoring Solution in **Downloads by Targets** • Injects Remote Monitoring Solution as **Software Updates** • Remotely installs Remote Monitoring Solution through **Websites visited by the Target**

Table 6: Feature-set of different FinFisher components (source Gama Group)

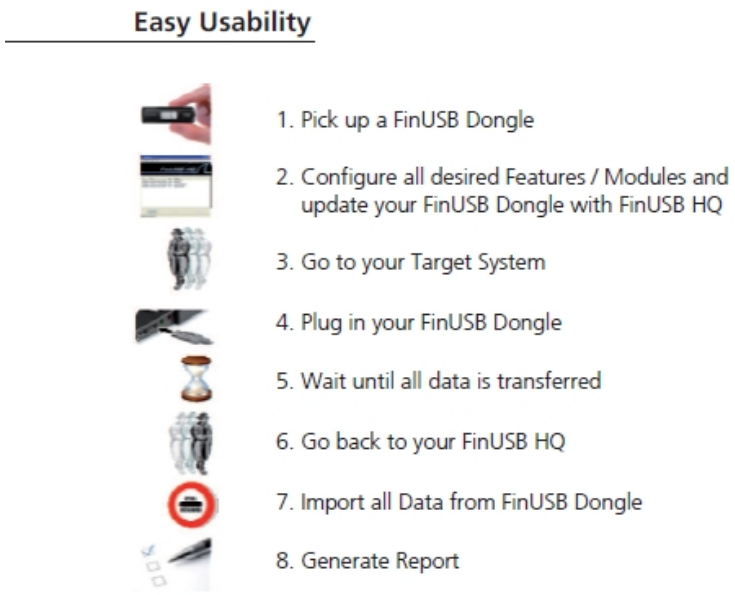

Easy Usability
1. Pick up a FinUSB Dongle
2. Configure all desired Features / Modules and update your FinUSB Dongle with FinUSB HQ
3. Go to your Target System
4. Plug in your FinUSB Dongle
5. Wait until all data is transferred
6. Go back to your FinUSB HQ
7. Import all Data from FinUSB Dongle
8. Generate Report

Figure 14: Steps for using FinUSB (source GammaGroup)

As indicated in the descriptions of some of the components (i.e. FinUSB Suite, FinFly USB, see Figure 14) they do not require IT-trained personnel for being operated. Other parts of the surveillance suite do obviously require the operating personnel to have solid knowledge about where or what to look for in a targeted system and for interpreting the information and data they retrieve from target systems.

The operation of the proper FinFisher components does, however, seem to be comparable to the operation of most of the mainstream software, judging by the screenshots of their user interfaces that are provided in the product portfolio. The FinSpy user manual[319] gives detailed instructions on how to install and operate the software. It reads pretty much like an installation manual of any other software package on the market and does definitely not require extensive IT knowledge to understand and follow its instructions. The user interfaces of the different components include options for multiple data views, structural analysis, definition of importance levels, filters and colour codes for facilitating the processing of the intercepted data.

Pre-defined and user adaptable reports are another feature of many of the FinFisher components. The FinUSB Suite, for instance, is providing the following list of reports, according to the information documented in the product portfolio:

- Generic
  - Generic Information
- Password
  - Windows Account Hashes
  - E-Mail Accounts
  - Messenger Accounts
  - Google Chrome Passwords
  - Firefox Passwords
  - Network Passwords
  - Protected Storage
  - Internet Explorer Accounts

[319] https://wikileaks.org/spyfiles4/documents/FinSpy-3.10-User-Manual.docx

- System
  - Windows Product Keys
  - Windows Updates
  - LSA Secrets
  - Current Processes
- Network
  - Network Adapters
  - Network Ports
  - Internet Explorer History
  - Mozilla Firefox History
  - Wireless Keys
  - Mozilla Firefox Cookies

The Remote Control System is another "*hacking suite for governmental interception*", sold by Hacking Team[320], an Italian security company. Target systems need to be infected via Email attachments or via drive-by installation of a malicious java applet[321]. The malware allows its controllers to monitor mouse coordinates, several instant messenger apps, location, internal webcam, clipboard contents, key strokes, running applications, web URLs, screen shots, internal mic, calendar data and alerts, device information and address book contents.

Regarding the **surveillance technology for Internet monitoring**, it includes IP interception technology which is intended to be operated by technical personnel as it is described in the installation guide of Blue Coat ProxySG series[322].

Internet interception technology passively intercepts communications from multiple interfaces[323] (e.g. 10/100 Mbit/s Ethernet with 1, 2 or 4 ports, Gigabit Ethernet 1000 Mbit/s (Fiber), T1 for 2 or 4 full duplex (FDX) connections, V.35 for 2 or 4 full duplex (FDX) connections, etc.) and data link-layer (i.e. OSI layer 2) protocols (e.g. Frame Relay, HDLC, CISCO HDLC, PPP, Bay PPP, MLPPP, 802.3/VLAN).

IP interception tools collect IP packets from the communication line where the taps or network probes have been located.

[320] http://www.hackingteam.it/

[321] http://www.cso.com.au/article/431882/_crisis_os_x_trojan_made_by_lawful_intercept_vendor_hackingteam/

[322] https://www.privacyinternational.org/sii/document/815764-116-blue-coat-instruction-installation-guide

[323] https://www.privacyinternational.org/sii/document/441689-poseidon-ip-information

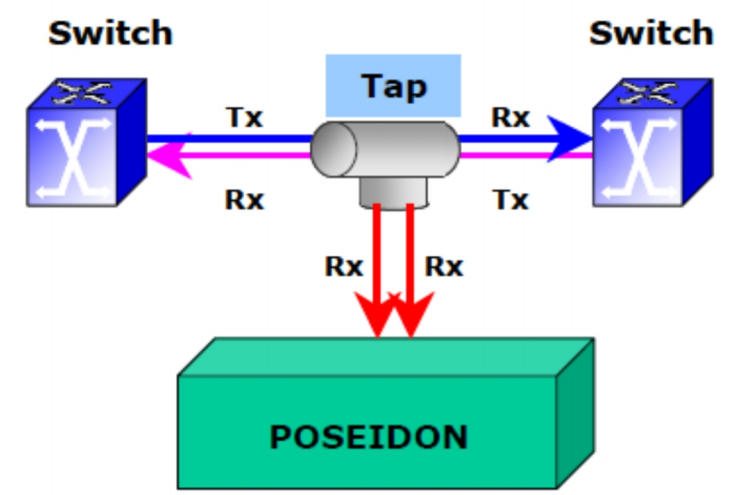


Figure 15: ELAMAN's POSEIDON Internet Monitoring Centre connected with a tap to the network (source Privacy International[323])

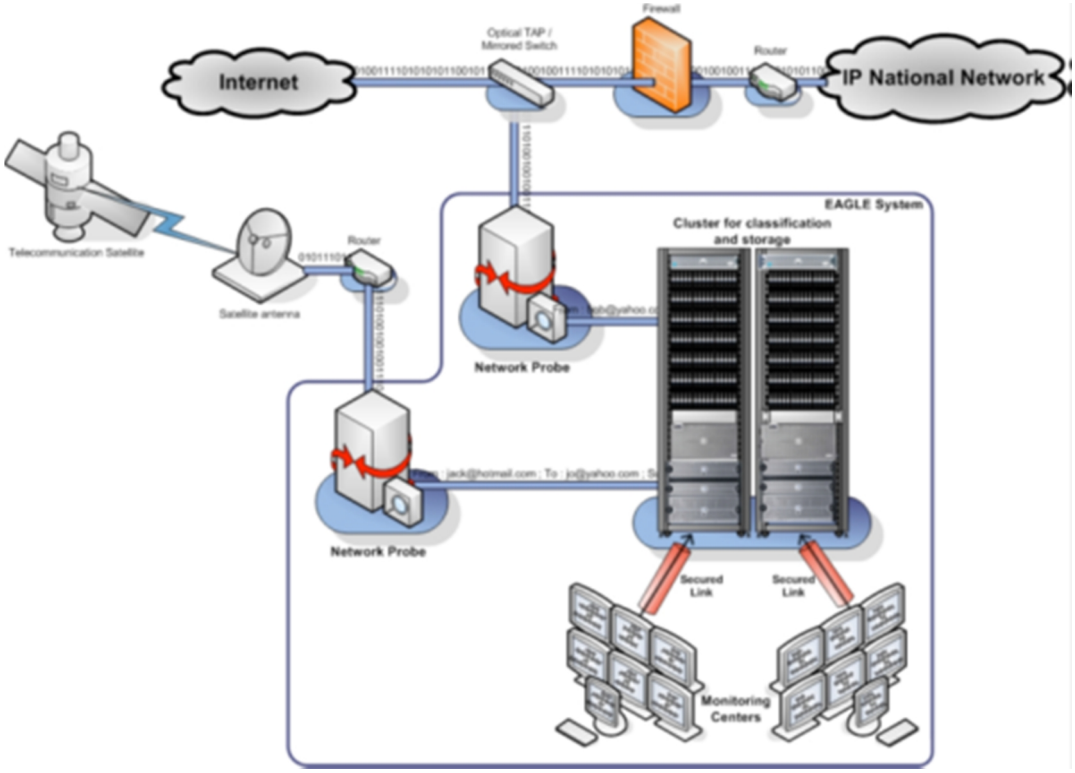


Figure 16: AMESYS' EAGLE System (source Privacy International[324])

This kind of technology can filter IP-based data prior to or after storing this data. If the incoming data is filtered prior to storing it, only the data matching the filter criteria will be stored.

Once the IP packets are stored in one or many databases, which are part of the Internet monitoring tool, the next step in the Internet monitoring process is to reconstruct the recorded IP-data. The reconstruction is protocol dependant and can be at any OSI level; e.g. at data link level Ethernet packets reconstruction, at network level IP packets reconstruction, at application level several protocols can be reconstructed such as SMTP, POP3, HTTP, IMAP4, Telnet, FTP, VoIP, etc.

[324] https://www.privacyinternational.org/sii/document/409210-99-amesys-eagle-glint-operator-manual

Although the IP interception tools are intended to be installed, configured and operated by technical personnel, the Graphical User Interface (GUI) for the analysis of the intercepted data can be used by less technical persons.

For instance, AMESYS' EAGLE system's *Man-Machine Interface* is a user-friendly web-based interface that includes buttons for searching and filtering and a screen that shows search and filtering results[324]. In addition, The GUI displays the status messages of the performed actions (such as searches) with different colours in order to facilitate their understanding by the user.

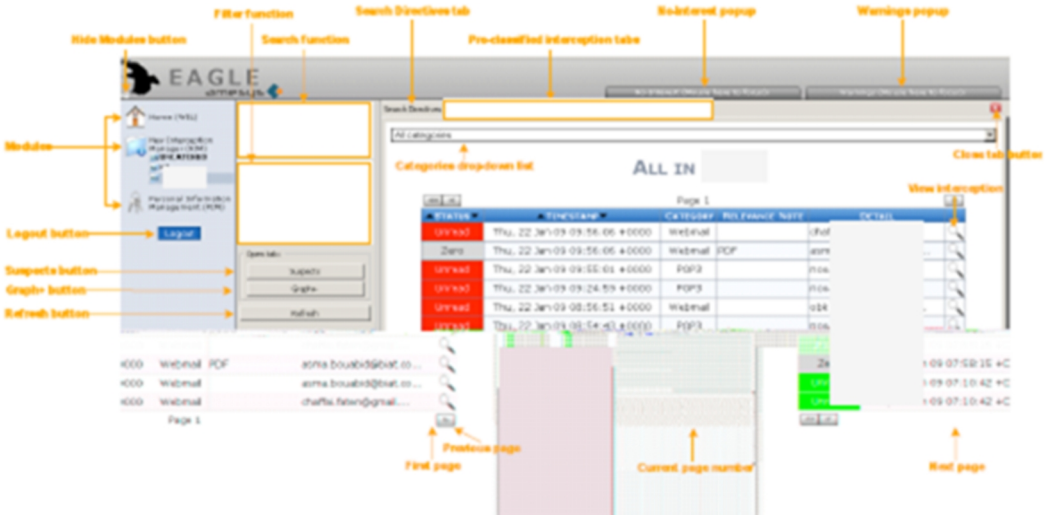


Figure 17: AMESYS' EAGLE System's Man-Machine Interface (source Privacy International[324])

As a result of filtering and search requests, the Internet monitoring tools show detailed list of recorded and reconstructed IP-based data such as HTTP sessions, emails, FTP sessions, chat sessions and VoIP sessions[308].

Figure 18: Screenshots of email, chat and voip sessions reconstructed with ELAMAN POSEIDON (Source Privacy international[323])

3.3 Question 21

What types of computing resources and power (CPU, Memory,Storage) are needed to monitor and analyse meta-data of an entire country over a period of 5 years?
What are the practical limits of what can be achieved with the current technology?
How is technology progressing in this domain, what kind of products will be on the market tomorrow compared to what is already available today ?

It is not possible to provide concrete metrics in terms of computing power or storage capacities necessary for monitoring and analysing metadata, because such a calculation would highly depend upon variables like the type of metadata, the population of and the penetration of internet and cell phones in the respective country, among others. Only relative numbers can be reported.

According to Cisco's Global Mobile Data Traffic Forecast of 2013 the current annual global mobile traffic amounts to 31,2 Exabytes (EB)[325] and will increase up to 190 EB in 2018[326]. The global yearly IP traffic currently adds up to 744 EB and is predicted to build up to 1584 EB in 2018[327]. Data provided by the BBC suggests that the overall internet traffic in 2011 was already reaching 330 EB[328]. No matter which estimation is closer to reality, metadata takes up only a fraction of that size.

Qosmos, a French company, suggests a 1:150 ratio between raw (intercepted) traffic and (filtered) metadata in the case of an email intercepted from a webmail page, in a product presentation of their DPI/NI tool[329]. Although this ratio is the result of an intelligent filtering of the raw data and may not be extrapolated to all types of intercepted data it gives an indication of the relation between the size of transmitted raw data and its corresponding metadata.

Much has been written in the media about the data centre the NSA is about to complete building in Utah and about the purpose and capacities of this facility. The NSA itself does not comment or publish any details on the project, which has led to speculations that the reason behind this datacentre is the objective of archiving (part or the entire) internet traffic data during a limited period, so as to be able to inspect it retrospectively. The estimations of experts on the projected storage capacities of this data centre fluctuate between 12.000 picabytes and various zetabytes or even yotabytes[330]. Since it is not possible to precisely determine the final storage capacity (which is classified information of the NSA) and despite the fact that this is only one among various data centres the NSA maintains[331], it is not possible to make a qualified statement on the *current* interception and storage capacities of the NSA, nor that of other national intelligence agencies. This does however not question national intelligence agencies' general ability of intercepting and storing IP and mobile data:

Based on the documents leaked by Snowden, NSA's Boundless Informant big data analysis and data visualization tool is capable of processing more than 97.000 million internet data records and 124.000 million telephony data records during a 30-day period. These data records only include metadata[332].

[325] 1 Exabyte = 1000 petabytes = 1 million terabytes = 1 billion gigabytes.
[326] http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf
[327] http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.pdf
[328] http://www.bbc.com/future/story/20130621-byte-sized-guide-to-data-storage
[329] https://www.wikileaks.org/spyfiles/files/0/77_201110-ISS-IAD-T6-QOSMOS.pdf
[330] http://www.forbes.com/sites/kashmirhill/2013/07/24/blueprints-of-nsa-data-center-in-utah-suggest-its-storage-capacity-is-less-impressive-than-thought/
[331] http://www.washingtonpost.com/world/national-security/nsa-growth-fueled-by-need-to-target-terrorists/2013/07/21/24c93cf4-f0b1-11e2-bed3-b9b6fe264871_story.html
[332] http://www.theguardian.com/world/interactive/2013/jun/08/boundless-informant-nsa-full-text

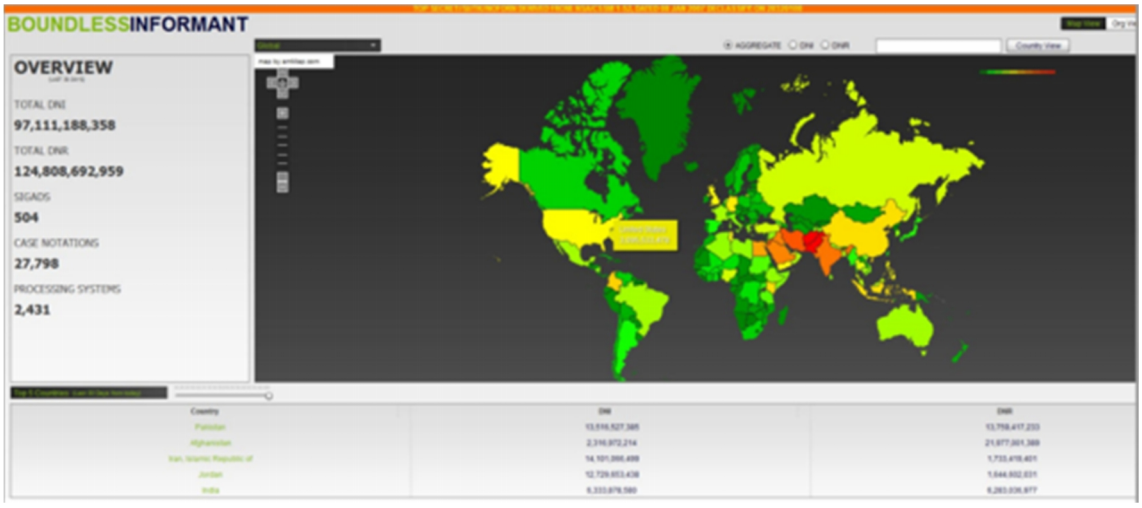


Figure 19: Screenshot of BOUNDLESS INFORMANT[333]

Figure 20 shows the monitoring capabilities in Germany for the period between 10 December 2012 and 8 January 2013[334]: more than 552 million telephony and internet data.

Figure 20: Communication monitoring capabilities in Germany

Figure 21 shows the monitoring capabilities in Spain and Italy in the same period but only for telephony data records[335]; 60 million telephony metadata in Spain and 46 million telephony metadata in Italy.

[333] Note: DNI means Digital Network Intelligence (i.e. data that is collected from the Internet) and DNR means Dialed Number Recognition (i.e. data that is collected from tracking phone numbers).
[334] http://www.spiegel.de/fotostrecke/photo-gallery-nsa-documentation-of-spying-in-germany-fotostrecke-99672-2.html
[335] http://www.spiegel.de/fotostrecke/photo-gallery-nsa-documentation-of-spying-in-germany-fotostrecke-99672-5.html

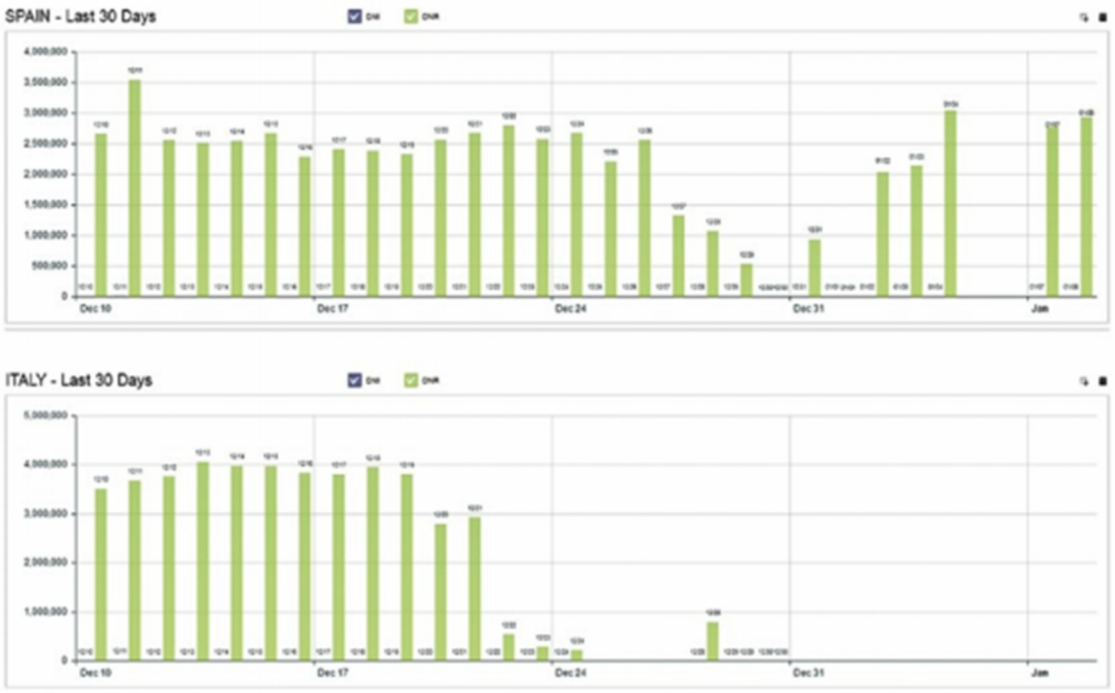


Figure 21: Monitoring capacity of telephone metadata in Spain and Italy

Other leaked documents show the volume of telephony metadata gathered worldwide that FASCIA, NSA's huge database which contains DNR (telephony metadata), can manage daily: almost 5.000 million records[336].

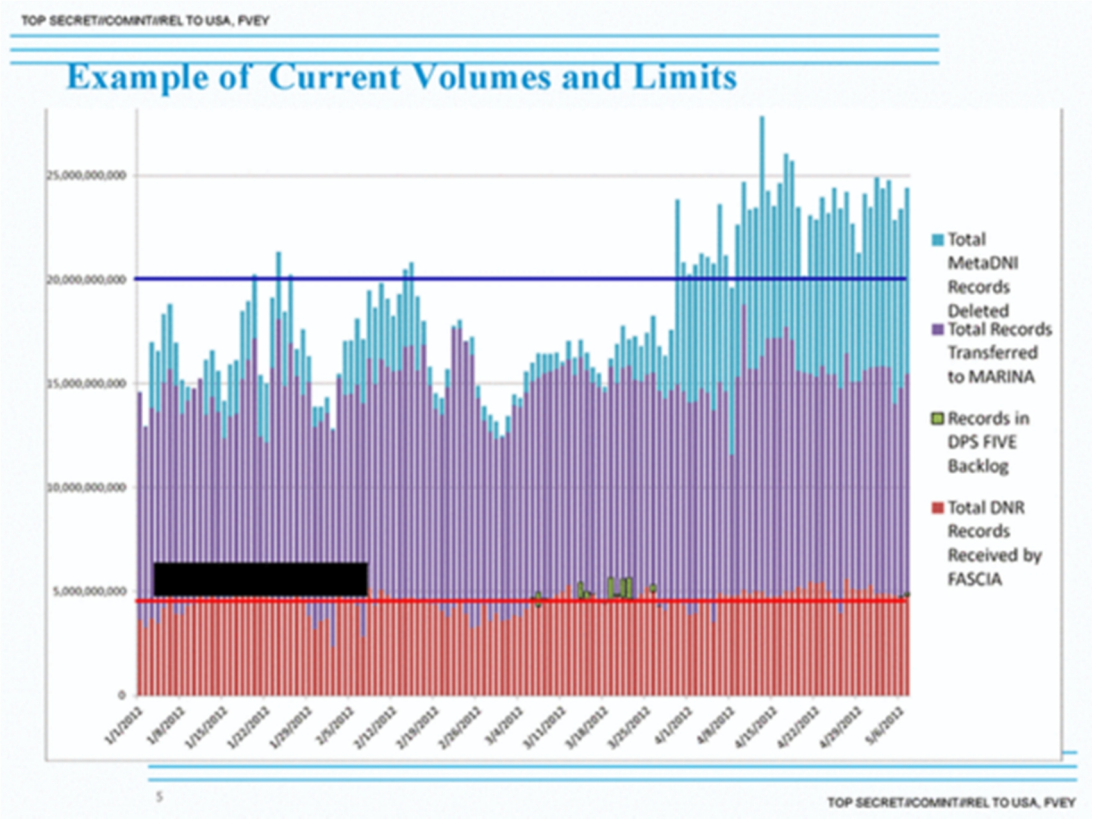

TOP SECRET//COMINT//REL TO USA, FVEY
Example of Current Volumes and Limits
Total MetaDNI Records Deleted
Total Records Transferred to MARINA
Records in DPS FIVE Backlog
Total DNR Records Received by FASCIA
TOP SECRET//COMINT//REL TO USA, FVEY

Figure 22: Current volume and limits of FASCIAs telephony metadata storage capacity

[336] http://apps.washingtonpost.com/g/page/world/what-is-fascia/637/#document/p1/a135288

Experts predict a continuous increase in storage capacities for the future, due to new solid state storage technologies and their combination with distributed cloud storage architectures. Already today, companies like Cleversafe[337] (among others like Facebook, Google, Amazon, etc.) do manage (distributed) database systems of over 10 EB storage capacity. This suggests that providing adequate storage capacities for intercepting large amounts of data (even in its raw format) will not be a technical or organisational problem in the future and that the fear of intelligence agencies being (technically) able to "*pull a complete backup of all internet traffic in real-time*" are not exaggerated at all.

But this does not mean that this information would be immediately accessible. Only the un-encrypted data could be directly analysed, while encrypted data would need to be decrypted first. The effort for deciphering encrypted data may potentially be huge, depending on the relative amount of such scrambled data and the strength of the encryption keys (assuming their intact integrity) and technologies used. This computing effort, combined with the CPU capacity needed to perform the complex algorithms needed for analysing these huge amounts of (Big) data is what is generating the bottleneck in today's endeavour to use such a strategy for surveillance purposes. At the same time this compelling need for more (and at best unlimited) processing power is what drives law enforcement and intelligence agencies to invest in researching graphene and quantum technologies[338].

3.4 Question 22

Are there concrete solutions currently available to help EU citizens to protect their metadata from organisations doing mass surveillance using commercial tools available on the market?
To which extent, can cryptography and/or anonymization help or not to prevent the collection and analysis of meta-data on a large scale by organisations doing mass surveillance?
Does the level of protection achievable by the end-user make any difference depending on whether the telecom operators, Internet service providers, or mobile service operators are providing (or not) their technical support to the organisations doing the mass surveillance using such commercial tools?
Is it possible for end-users to protect themselves irrespectively of any technical complicity/cooperation between network operators with mass surveillance organisations running such commercial tools?

Please refer to Question 6 for examples of concrete technologies and solutions available in the market that the citizens can use in order to protect their privacy on-line and to Question 7 for mechanisms that help to prevent cryptography problems

As can be seen from the analysis of such solutions, the degree of safety from metadata tracking that these solutions achieve depends on the type of encryption and the communication layer in which the encryption is applied. Most of the solutions are valuable for keeping content confidentiality and integrity, while metadata confidentiality and integrity depends on which transport layers the encryption is applied.

The more layers in which the encryption is applied the better the protection, as more metadata can be hidden from third parties' eyes. While some of the tools like the use of IPSec and TOR (TLS based network tunnelling) provide greater protection as they obfuscate most of the metadata (even the network layer metadata, e.g. IP addresses of the source and the destination), others do only hide

[337] http://www.cleversafe.com/news-reviews/cleversafe-press-releases/2012-press-releases/cleversafe-announces-10-exabyte-storage-system-configuration
[338] See also Question 32

metadata at higher layers and do leave lower layers' metadata unprotected. *For example, even if using a securely encrypted end-to-end voice service such as Silent Circle, your ISP (and anyone else watching) will still be able to tell who you are making the call to, whether you are calling from your usual location, how often you call, how long you usually chat, and much more. With such information it would be trivially easy to discover that you were having an affair (for example)*[339].

Experts recommend using combined solutions that encrypt metadata both at application layer and network layer. The metadata in data link and physical layers are the most difficult to protect by the citizen himself. Protecting metadata in these layers is mostly in the hands of network providers and device vendors.

When it comes to the use of cryptographic protocols (HTTPS, IPSec, etc.), the level of protection achievable is also dependent on the accuracy in which the solutions do implement the protocols. The accuracy is a factor which may be influenced by mass surveillance organisations collaborating in or conditioning such implementations. The protection degree thus depends on which tools are used and the level of cooperation between network operators with mass surveillance organisations.

3.5 Question 23

What are the tools commercially available for breaking cryptography/encryption for mass surveillance purposes, - either by exploiting compromised cryptographic information (ie, encryption keys) that should normally have been kept confidential, but that have been made available to mass surveillance organisations to allow them to decrypt the traffic they intercepted, - or by using hacking exploits, security vulnerabilities, and backdoors. Can these commercial tools be used to access to un-encrypted content on-line real-time during the exchange of information as it happens, or off-line after the exchange has taken place pending proper off-line decryption? How efficient are these tools at decrypting messages if the "encryption keys" are unknown?

According to Privacy International, in ISS (Intelligence Support Systems for lawful interception, electronic surveillance and cyber intelligence gathering) World[340] Middle East 2014 conference TeleStrategies[341] proposed 7 different means for defeating encryption[342], shown in Figure 23.

[339] https://www.bestvpn.com/the-ultimate-privacy-guide/#meta
[340] http://www.issworldtraining.com/
[341] http://www.telestrategies.com/
[342] https://www.documentcloud.org/documents/1215458-1299-telestrategies-presentation-challenges.html#document/p46/a178126

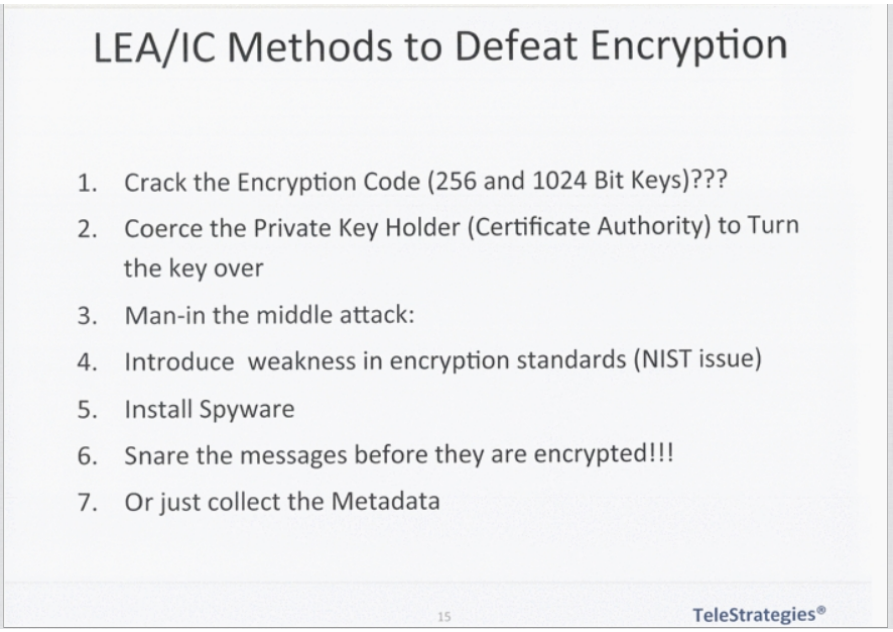

LEA/IC Methods to Defeat Encryption
1. Crack the Encryption Code (256 and 1024 Bit Keys)???
2. Coerce the Private Key Holder (Certificate Authority) to Turn the key over
3. Man-in the middle attack:
4. Introduce weakness in encryption standards (NIST issue)
5. Install Spyware
6. Snare the messages before they are encrypted!!!
7. Or just collect the Metadata
15
TeleStrategies®

Figure 23: 7 means for defeating encryption (source TeleStrategies)

The Spy Files[343] catalogue published by Wiki Leaks lists a number of examples of commercial tools that are able to listen to communications even if they are encrypted. One of them is Hacking Teams' Remote Control System (RCS) Trojan, which is "*a stealth system for attacking, infecting and monitoring computers and smartphone*"' for targeted surveillance. Hacking Teams' RCS[344] would therefore pertain to the category number 5 in the means proposed by TeleStrategies[345].

SSL Locksmith by ACCESSDATA[346] is defined as a solution which *works with any packet analysis or capture solution, such as IDS, IPS, DLP, network forensics and web content monitoring solutions, to reveal the content of encrypted network communications - dramatically improving an organization's ability to monitor encrypted traffic for data leakage and ensure compliance with corporate and government-mandated regulations.*

Some National intelligence agencies are also suspect of having developed such kind of tools for undermining the encryption in secure communications. The best examples of tools of the like are those that, according to The Guardian on Snowden revelations[347], were developed by NSA and by GCHQ. The program codenamed Bullrun served to break the encryption in widely used online protocols, such as HTTPS, voice-over-IP and Secure Sockets Layer (SSL), which are used to secure online sensitive transactions.

The GCHQ would also have developed the British version, codenamed Edgehill. "*Edgehill's initial aim was to decode the encrypted traffic certified by three major (unnamed) internet companies and 30 types of Virtual Private Network (VPN) – used by businesses to provide secure remote access to their systems. By 2015, GCHQ hoped to have cracked the codes used by 15 major internet companies, and 300 VPNs.*" The same report by the Guardian cites another program codenamed Cheese Name that served to select those "enc*ryption keys that might be vulnerable to being cracked by GCHQ supercomputers*"[348].

The Washington Post[349] documented the $79.7 million research program *Penetrating Hard Targets project* of the NSA which included the research on the future development of "*a cryptologically useful*

[343] https://wikileaks.org/the-spyfiles.html
[344] See also Question 19
[345] http://www.telestrategies.com/
[346] http://accessdata.com/solutions/cybersecurity/ssl-locksmith
[347] http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security
[348] http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security
[349] http://apps.washingtonpost.com/g/page/world/a-description-of-the-penetrating-hard-targets-project/691/

quantum computer" which would allow NSA "*breaking all public key encryption, including RSA, which is used for most secure websites and for encrypted e-mail conversations*." Such a tool would allow cracking most types of encryption[350] by brute-force without the need of knowing the encryption keys.

Due to the lack of (access to) clear evidences, it is difficult to ensure whether these tools were only aimed at lawful targeted surveillance or also envisaged for mass surveillance purposes. In order to use them in a massive scale they would need to be combined with powerful storage and processing capacities, which are not difficult to achieve with current computer technologies[351]. The achievable scale would therefore depend mostly on the available resources in monetary terms.

Nevertheless, other approaches to defeating communications encryption seem to be more prone to mass surveillance, for example, introducing weaknesses in encryption standards (means number 4 in Figure 23) or collaborate with commercial vendors of security tools.

One of the computer science courses filed in Professor Eric Roberts' website at the Stanford University[352] reports that in November 2007 the NSA was suspected of inducing backdoors in the standards for random number generators that are used for encryption[353]. "*NIST Special Publication 800-90, NIST's 2007 official standard for random number generators, is believed to have included a secret backdoor on the behalf of the NSA.*"

The National Institute of Standards and Technology (NIST) is the responsible of specifying which techniques are approved for use in the US public and private sectors. The importance of compromising the random-number generators lies in the fact that these random numbers (or seeds) are used in cryptography for the generation of the encryption keys that are needed to cipher and decipher the encrypted information. "*If an agent is able to compromise the random-number generator, that agent has most likely compromised any encryption scheme using that random-number generator.*"'[354]

The course documentation explains: "*In cryptography specifically, a backdoor would allow an intruder to access the encrypted information without having the correct credentials. The backdoor would either a) allow the intruder to guess the access key based on the context of the message or b) allow the intruder to present a skeleton key that will always grant him access.*"

Other examples of suspected government meddling in encryption related standards can be found. A comprehensive report on "Technology, Policy, Law, and Ethics Regarding U.S. Acquisition and Use of Cyberattack Capabilities"[355] by the US National Academy of Science, says "*In addition, it is entirely possible that certain technical problems have solutions that are today classified and thus not available to the world at large. In the domain of cryptography, it is known that the British Government Communications Headquarters (GCHQ; the UK equivalent of the National Security Agency) knew of public key encryption and in particular of the RSA algorithm for public key encryption several years before they were announced in the open literature*[356]*. Thus, one might reasonably presume that there may well be technical approaches to various forms of cyberattacks that are known, at present, only on the "inside"*".

With respect to collaboration with cryptography breaking tool vendors, in the top secret NSA's "Classification guide for cryptanalysis" revealed by Snowden[357] it is documented that the NSA

[350] http://www.washingtonpost.com/world/national-security/nsa-seeks-to-build-quantum-computer-that-could-crack-most-types-of-encryption/2014/01/02/8fff297e-7195-11e3-8def-a33011492df2_story.html
[351] See also Question 21
[352] http://cs.stanford.edu/people/eroberts/cs201/projects/ethics-of-surveillance/tech_encryptionbackdoors.html
[353] https://www.schneier.com/essays/archives/2007/11/did_nsa_put_a_secret.html
[354] http://cs.stanford.edu/people/eroberts/cs201/projects/ethics-of-surveillance/tech_encryptionbackdoors.html
[355] http://www3.nd.edu/~cpence/eewt/Owens2009.pdf
[356] Peter Wayner, "British Document Outlines Early Encryption Discovery," New York Times, December 24, 1997, available at http://www.nytimes.com/library/cyber/week/122497encrypt.html#1.
[357] http://www.theguardian.com/world/interactive/2013/sep/05/nsa-classification-guide-cryptanalysis

seemed to have collaborated with commercial organisations to get insights on the cryptographic details of their commercial security systems and products. This would have left exposed to NSA eyes not only targeted encrypted communications but those of anyone using such commercial services.

After Snowden's revelations and leaked documents[358], The Guardian[348] reported that "*the National Security Agency and its UK counterpart GCHQ have broadly compromised the guarantees that internet companies have given consumers to reassure them that their communications, online banking and medical records would be indecipherable to criminals or governments.*" And stated the agencies have used a number of methods to try to overcome ubiquitous encryption in Internet communications. "*Those methods include covert measures to ensure NSA control over setting of international encryption standards, the use of supercomputers to break encryption with "brute force", and – the most closely guarded secret of all – collaboration with technology companies and internet service providers themselves. Through these covert partnerships, the agencies have inserted secret vulnerabilities – known as backdoors or trapdoors – into commercial encryption software.*"

3.6 Question 24

Is the sale of such commercial products and services, which are openly advertised as allowing the mass surveillance of users of an entire country, regulated in Europe?
Are these commercial companies operating legally in their home European jurisdiction?

Despite there are many products and services that serve for mass surveillance (see Question 19), few are commercial solutions openly advertised for mass surveillance purposes.

Verint[359], a multinational company headquartered in Melville, NY, is one of those few that speaks about "mass" and "nationwide" scale when presenting their interception solutions, as can be observed in Figure 24. In their words, "*Unlike target interception solutions, which intercept only specific communications, mass interception solutions intercept all communication traffic.* ***These solutions allow agencies to investigate criminal and terrorist events after they happen*** *to help prevent escalations and new occurrences.* ***Through retrospective analysis and investigation of mass communications****, agencies can also identify unknown or potential threats and gather new leads. However, agencies are challenged to distill the most timely and important intelligence from the huge volumes of communications captured.*"

[358] American Civil Liberties Union – The NSA Archive https://www.aclu.org/nsa-documents-search
[359] http://www.verint.com/index

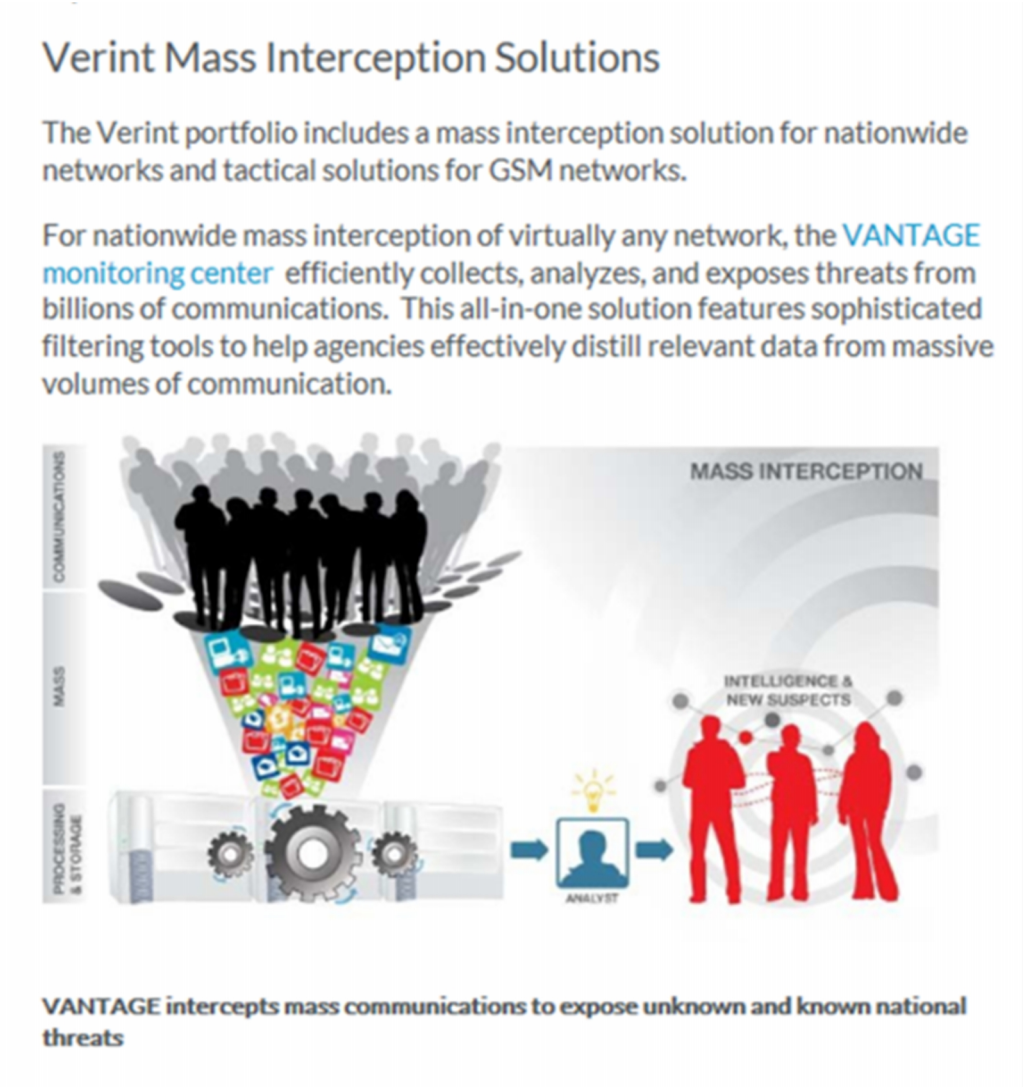
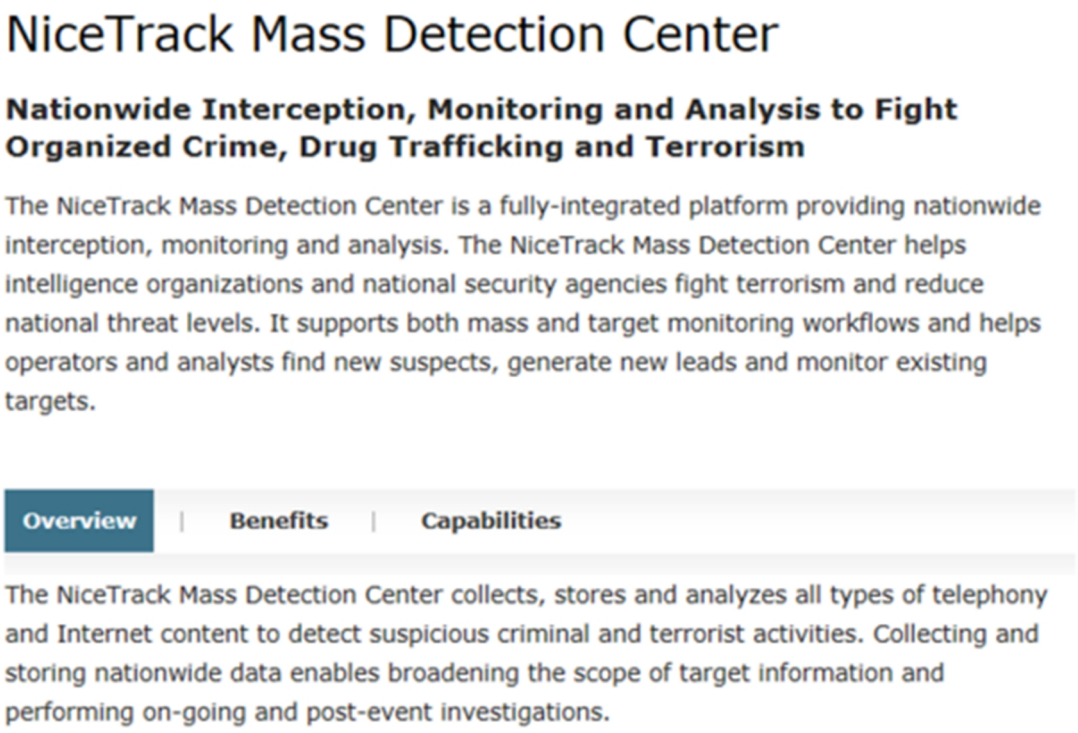

Verint Mass Interception Solutions
The Verint portfolio includes a mass interception solution for nationwide networks and tactical solutions for GSM networks.
For nationwide mass interception of virtually any network, the VANTAGE monitoring center efficiently collects, analyzes, and exposes threats from billions of communications. This all-in-one solution features sophisticated filtering tools to help agencies effectively distill relevant data from massive volumes of communication.
MASS INTERCEPTION
COMMUNICATIONS
MASS
PROCESSING & STORAGE
INTELLIGENCE & NEW SUSPECTS
ANALYST
VANTAGE intercepts mass communications to expose unknown and known national threats

Figure 24: Verint marketing information

Nice[360] is another leading multinational that does also market their products openly for mass surveillance purposes.

[360] http://www.nice.com/



NiceTrack Mass Detection Center

Nationwide Interception, Monitoring and Analysis to Fight Organized Crime, Drug Trafficking and Terrorism

The NiceTrack Mass Detection Center is a fully-integrated platform providing nationwide interception, monitoring and analysis. The NiceTrack Mass Detection Center helps intelligence organizations and national security agencies fight terrorism and reduce national threat levels. It supports both mass and target monitoring workflows and helps operators and analysts find new suspects, generate new leads and monitor existing targets.

Overview | **Benefits** | **Capabilities**

The NiceTrack Mass Detection Center collects, stores and analyzes all types of telephony and Internet content to detect suspicious criminal and terrorist activities. Collecting and storing nationwide data enables broadening the scope of target information and performing on-going and post-event investigations.

Figure 25: Nice marketing information

Yet another example of solutions that advertise their capacities for mass surveillance is the Eagle product by the French company Amesys, although its features are not publicly available, just as with most of the mass surveillance tools'. Presumably confidential brochures can though be found e.g. in documentcloud.com[361] and in es.scribd.com[362].

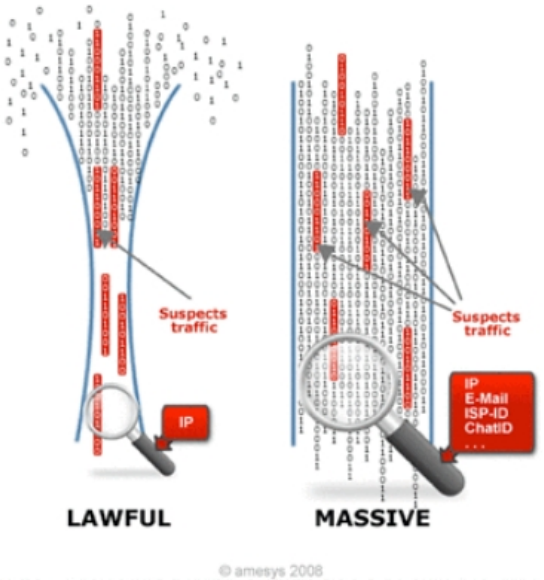


Figure 26: Eagle product brochure differentiating between lawful and mass interception

[361] https://www.documentcloud.org/documents/409136-21-200810-iss-prg-amesys.html#document/p9/a119301

[362] http://es.scribd.com/doc/73812289/Brochure-commerciale-du-systeme-de-surveillance-Eagle-d-Amesys-Bull

The OWNI.EU information website declares that "*Amesys surveillance system is used to intercept and analyze the entirety of the telecommunications network, to the scale of an entire country*" [363] and Privacy International comments on the slide shown in Figure 26 "*Amesys' Eagle system makes the particular distinction between the two focuses of it's system. The first is Lawful Interception which presumes a legally based framework in which to conduct surveillance, targeting specific suspects and avoiding interception of other content. The other option is Mass, looking at everyone's information as it moves through the communication framework and picking out the information relevant to you.* ***It also implies that there is no legal framework for this type of surveillance*** *either considering the former option*"

These companies do not usually provide publicly detailed information on their services or products, and often their websites are merely one or two page sites that succinctly proclaim the organization's skills and capabilities, including mass surveillance, but not much more. See the example of Clear Trail[364] (an Indian company according to Buggedplanet[365]) in Figure 27. The only information on their website is shown in this screen, although some information on their products for *Mass monitoring of IP and Voice networks* is archived by WikiLeaks in the Spy Files[366].

ClearTrail
Welcome to CLEARTRAIL TECHNOLOGIES
For a decade, ClearTrail has been an established consultant and solution provider for Law Enforcement and Intelligence Agencies. ClearTrail solutions enable the LEAs to perform MASS, TARGETED & TACTICAL Monitoring and Analysis across a variety of communication networks. ClearTrail's Research & Development centers are committed to build path breaking lawful interception solutions for LEAs.
For more information reach us at info@clear-trail.com
To Start an exciting career with us, mail your resume at jobs@clear-trail.com
CELEBRATING
10
Years

Figure 27: ClearTrail corporate information

3.7 Question 25

> ***It seems, according to the press, that the products of some of these companies have been used by foreign governments (non EU) for illicit political intelligence purposes. What control these companies have over the possible misuse of their products by their customers and to which extend are they legally responsible for it and the consequences of the misuse of their products outside Europe?***

Answers to this question have been elaborated under Question 19.

[363] http://owni.eu/2011/09/12/a-guide-to-libyas-surveillance-network/
[364] http://www.clear-trail.com/
[365] http://buggedplanet.info/index.php?title=CLEARTRAIL
[366] https://www.wikileaks.org/spyfiles/docs/cleartrail/111_tactical-interception.html

3.8 Question 26

Are there possible commercial conflicts of interest, such as:
- Commercialisation of security solutions on one hand, and commercialisation of mass surveillance products on the other hand, inducing the risk of some commercial backdoors implemented on purposes to stimulate sale of mass surveillance products relying on these backdoors?
- Commercialisation of security products and services to both national security agencies and to sensible organisations monitored by the same national security agencies, hence inducing the risk that some companies might be tempted to accept offers from their main governmental customers to implement backdoors in their products.

As a matter of fact the commercialisation of security solutions and mass surveillance programs is complementary in many cases. The example of BlueCoat, a company that is specialized in online security but it is well known in the market surveillance for its DPI technology based equipment[280]**Error! Bookmark not defined.**, clearly shows that **the same vendor develops and markets solutions on both segments**.

However, it can only be speculated whether there exists a commercial conflict of interest caused by the creation of backdoors in security tools, in order to stimulate the sale of mass surveillance tools that work on these backdoors. **The authors of this study did not find any evidence or reports that would confirm such practices**.

On the other hand, it seems at first sight that the scenario of collaborations between vendors of security products and national security agencies does exist and present conflicts of interests. Based on the documents leaked by Snowden, **there are evidences that the NSA has compromised a wide range of systems** such as network security equipment, PCs, and hard drives by including backdoors[367][368][369]. These backdoors enable national security agencies, the NSA in this case, to intercept internet traffic (both data and metadata) in the case of backdoors included in network equipment and to access data stored in devices such PCs and hard drives.

However, there are no evidences that the security product vendors have cooperated with the NSA to install the backdoors[369]. In fact, all affected vendors claimed they were unaware of the vulnerabilities their systems could have or of any modification that could have been done to these systems[367]. According to recent statements by Snowden during an interview in "The Nation", the companies did not know about the backdoors: "***Companies did not know it****. They said, "Well, we gave the NSA the front door; we gave you the PRISM program. You could get anything you wanted from our companies anyway—all you had to do was ask us and we're gonna give it to you." So* ***the companies couldn't have imagined that the intelligence communities would break in the back door, too—but they did, because they didn't have to deal with the same legal process as when they went through the front door.***"[370].

[367] http://www.spiegel.de/international/world/catalog-reveals-nsa-has-back-doors-for-numerous-devices-a-940994.html
[368] https://gigaom.com/2013/12/29/nsas-backdoor-catalog-exposed-targets-include-juniper-cisco-samsung-and-huawei/
[369] http://www.infoworld.com/article/2609310/hacking/apple--cisco--dell-unhappy-over-alleged-nsa-back-doors-in-their-gear.html
[370] http://www.thenation.com/article/186129/snowden-exile-exclusive-interview#

Thus, if the security products vendors have no knowledge of modifications in their systems and devices, there can't be a conflict of interest at all, since the responsible of including these backdoors are the national security agencies, without the cooperation of security products vendors.

Besides backdoors implementation on ICT systems and devices, there are other powerful mechanisms that enable external agents to gain access on third party systems. For instance, zero-day vulnerability exploits allow exploiting a publically not discovered vulnerability in a system, software or device; i.e. a flaw in a system is discovered and the zero-day exploit consist of a tool or code that exploits this flaw before a fix or patch is available.

Surveillance vendors such as Gamma Group and Vupen Security currently sell this type of intrusion technology[371][372]. As described by Privacy International[373]: *By using the FinFly Exploit Portal, governments can deliver sophisticated intrusion technology, such as FinSpy, onto a target's computer. While it's been previously advertised that Gamma use fake software updates from some of the world's leading technology companies to deliver FinSpy onto a target's computer, the exploit portal puts even more power in the hands of government by offering more choices for deployment. Astonishingly, FinFly Exploit Portal guarantees users four viable exploits for some of the most-used software products in the world, such as Microsoft's Internet Explorer and Adobe's Acrobat programme.*

According to a report published by the Citizen Lab of the University of Toronto, **there is a professional alignment between exploit sellers and surveillance vendors**: "*While these actors are natural business partners, the conclusion we draw is that the marketplace for exploits and surveillance software, despite the opacity and competitiveness for government contracts, can also be cozy, with vendors regularly working together to sell products and solutions to clients. While this collaboration may offer a one-stop-shop experience for purchasers, it also helps tie vendors, campaigns, companies, and countries back together when investigated.*" [374]

3.9 Question 27

How can end-users detect that they are subject to break of cryptography/encryption for mass surveillance purposes?
What can end-users do to protect themselves against this threat?
What are example of appropriate "backdoor free" security solutions, encryption and anonymization commercial products available to seek optimal protection?
Who are the leaders in this domain and are European companies on the leading edge?
How easy or difficult is it in practice for end-users to use these "commercial products" to protect themselves against encryption breaks across different types of platforms including mobile devices?

As previously mentioned (see Question 26), apart from backdoors there are other powerful technological mechanisms to gain access on the systems and devices, such as exploits. The security and cryptography expert Bruce Schneier pointed out "*Your anti-virus software won't detect them, and you'd have trouble finding them even if you knew where to look*"[375]. Therefore, **average end users cannot detect that they are subject to break of cryptography/encryption**. Even skilled experts find

[371] https://www.techdirt.com/articles/20140311/07593926528/surveillance-security-companies-set-up-zero-day-exploit-portals-governments-to-use-offensive-actions.shtml
[372] http://www.vupen.com/english/services/lea-index.php
[373] https://www.privacyinternational.org/news/blog/exploiting-privacy-surveillance-companies-pushing-zero-day-exploits
[374] https://citizenlab.org/2014/02/mapping-hacking-teams-untraceable-spyware/
[375] http://www.theguardian.com/world/2013/sep/05/nsa-how-to-remain-secure-surveillance

difficulties to discover and trace encryption breaches, being clear examples the uncovering and detention of (professional) cyber criminals who were relying on anonimization and encryption technologies for disguising their activities.

However, security researchers from the Citizen Lab of the University of Toronto have been capable of detecting some traces related to two of the most powerful malware currently used for mass surveillance purposes: Gamma Group's FinFisher[281] and Hacking Team's Remote Control System[374]. Based on this work, **a free open-source tool named Detekt**[376] has been released in November 2014. This tool **scans Windows computers for traces of FinFisher and Remote Control Systems**. Detekt does not replace existing Antivirus, neither existing security products, but does complement them[377]. Detekt was developed by Claudio Guarnieri and is released in partnership with Amnesty International, Digitale Gesellschaft, Electronic Frontier Foundation and Privacy International.

Under Question 6 a collection of examples of concrete technologies and solutions available in the market are documented, which the citizens can use in order to protect their privacy on-line.

End-users have to utilize best practices and tools in order to protect their anonymity, privacy and confidentiality. Best practices include the use of specific tools that implement secure cryptography algorithms for protecting personal data storage and transmission, and the use of novel untraceable browser and anonymization networks such as TOR for preserving privacy. The capacity of the protocols and tools listed in Question 6 for the encryption of the content is not unlimited, but the degree of privacy of metadata tracking that these solutions can achieve depends on the type of encryption and the communication layer in which the encryption is applied (see Question 22 for more information on metadata encryption possibilities).

Many of the recommended tools for privacy protection are free and open-source initiatives such as GPG Tools[378], TOR, HTTPS Everywhere (part of the TOR ecosystem[379]) and TrueCrypt. Any developer around the world can join the community and contribute to the code of each of the tools. Even though TOR is an open-source initiative, it is registered as a non-profit organization in the USA which consists of many volunteers and few employees[380]. Other commercial tools are also located in USA such as Apple's FileVault[381], Microsoft's BitLocker[382], and Abine's Blur[383].

Some of the suggested tools are developed by Europe based companies such as Boxcryptor[384] (Germany), Cloudfogger[385] (Germany), StartPage[386] (The Netherlands), and StartMail (Germany).

In January 2014 the "Forbes" magazine published an article which highlights the leadership of German security start-ups on an existing global demand for online privacy protection[387]. The company ZenMate is distinguished as a privacy and security firm: *Launched in July 2013, the Berlin-based start-up provides end users with secure, encrypted access to any website, from anywhere, via a Virtual Private Network (VPN)-style connection. And in just over a year, it has grown its registered users to more than 5 million, including a 400% increase in the last six months alone. And they are by no means all based in Germany*[387].

[376] https://github.com/botherder/detekt/tree/v1.7
[377] https://resistsurveillance.org/intentions.html
[378] https://gpgtools.org/
[379] https://www.torproject.org/getinvolved/volunteer.html.en
[380] https://www.torproject.org/about/corepeople.html.en
[381] http://support.apple.com/en-us/HT4790
[382] http://windows.microsoft.com/en-us/windows7/products/features/bitlocker
[383] https://dnt.abine.com/#register
[384] https://www.boxcryptor.com/en/home
[385] http://www.cloudfogger.com/en/
[386] https://startpage.com/uk/company.html?
[387] http://www.forbes.com/sites/alisoncoleman/2014/10/01/germanys-security-start-ups-leading-a-global-demand-for-online-data-privacy/

3.10 Question 28

If end-users start to protect themselves extensively against mass surveillance threats using such privacy-enhancing tools, will law enforcement and/or national security agencies be able to continue to fight crime and terrorism or will they go "dark and blind" forever?

The extensive use of privacy-enhancing tools, most of all encryption tools, will certainly reduce law enforcement, national security and intelligence agencies' capacities of intercepting, processing and analysing massive amounts of data. This does, however, not reciprocally imply that they will go "dark and blind" forever, but that **they would need to shift towards a more focussed and targeted approach of crime prevention and investigation**[388].

The technical features described in the ANT catalogue[428,389] of hard and software interception tools available to NSA's TAO group give an impression of the level of **sophistication and technical excellence available to national security agencies** and that these **go far beyond those available to citizens** for "going stealth".

The recommendations for users to protect from mass surveillance all go into the same direction and put strong end-to-end encryption on top of the list of most powerful countermeasures. The announcement of Apple and Google to implement default encryption of all information in their new OSs has led the US Attorney General and the Director of the FBI to publically warn that this measure may leave law enforcement and/or national security agencies "in the blind" when trying to access evidences on phones equipped with these OS and that this reason will make the iPhone6 become "*the smartphone of choice for all paedophiles*".[390]

Edward Snowden, in an interview to "The Nation" newspaper early October 2014 rebutted that argument as follows: "*They were suggesting, "We have to be able to have lawful access to these devices with a warrant, but that is technically not possible on a secure device. The only way that is possible is if you compromise the security of the device by leaving a back door."* **[…]** *It is not possible to create a back door that is only accessible, for example, to the FBI. And even if it were, you run into the same problem with international commerce: if you create a device that is famous for compromised security and it has an American back door, nobody is* [going to] *buy it. Anyway,* ***it's not true that the authorities cannot access the content of the phone even if there is no back door. When I was at the NSA, we did this every single day****, even on Sundays. I believe that encryption is a civic responsibility, a civic duty*."[370]

3.11 Question 29

What are the technological and organisational options to prevent "Mass surveillance" while at the same time allowing competent national authorities to fight cyber-crime and protect "National Security Interests"?
To which extend are these objectives compatible or incompatible with each other from a scientific and technological perspective?

Technical options for preventing mass surveillance are described in Question 6 and essentially consist in hiding IP traces and encrypting content and communication channels. Every cybercriminal already makes use of these options in a very professional way and law enforcement and national security

[388] see also Question 29
[389] see also Question 32
[390] http://www.fbi.gov/news/speeches/going-dark-are-technology-privacy-and-public-safety-on-a-collision-course

agencies do have - and certainly need the right to have - the technical means to combat them. So it is rather an organisational question of how the usage of these technical capabilities by these national authorities can be limited to a legally sanctioned frame of action.

Such a legal frame of action would need to be internationally sanctioned and approved, since the global borderless distribution of internet and telecommunication infrastructure and business would vitiate such an agreement if only a few major countries did not agree to it. This scenario of divergent national interests was also mentioned in the Report of the Plenary Sitting of the Committee on Civil Liberties, Justice and Home Affairs in February 2014, which states that "*negotiations on the* [Framework] *agreement* [between EU-US on data protection in the field of police and judicial cooperation] *have not progressed because of the US Government's persistent position of refusing recognition of effective rights of administrative and judicial redress to EU citizens*" [391].

It is therefore not so much a question of trade-off between national security and privacy but rather one of finding the right equilibrium between these two social values. Caspar Bowden, independent privacy advocate, describes this equilibrium as the choice between "*the breaching of the principle against blanket collection which voids the essence of the right to private life, or a lawful regime of targeted and proportionate preservation, which nevertheless might result in a residuum of crimes that might have been prevented or detected under a blanket retention regime*"[392].

The European study about National programs for mass surveillance of personal data in EU MS and their compatibility with EU law clearly states that: "*an analysis of Europe's surveillance programs cannot be reduced only to the question of the proper balance between data protection and national security and to technical capabilities understood by experts. Rather, it has to be framed in terms of collective freedoms and the nature of democratic regimes*"[393].

[391] REPORT on the US NSA surveillance programme, surveillance bodies in various Member States and their impact on EU citizens' fundamental rights and on transatlantic cooperation in Justice and Home Affairs (2013/2188(INI))

[392] http://blog.privacystrategy.eu/public/published/Submission_ISC_7.2.2014_-_Caspar_Bowden.pdf

[393] http://www.europarl.europa.eu/RegData/etudes/etudes/join/2013/493032/IPOL-LIBE_ET%282013%29493032_EN.pdf

4 THEME 4 - TECHNICAL CREDIBILITY OF NATIONAL SECURITY AGENCIES HACKING CAPABILITIES

4.1 Question 30

Following the PRISM and other NSA spying scandals, there have been some allegations that some telecom operators were cooperating secretly with some national security agencies to allow them to tap and collect traffic flowing over their network infrastructures (against some form of monetary compensation). Are there credible informations leaks confirming that such deals took place, and have these leaks been publicly validated or invalidated by the organisations concerned?

A number of recent reports, most of which are based on material that was leaked by Edward Snowden, lead to the impression that intelligence agencies are using the state of the art of science and technology to its highest extend possible, in order to gather SIGINT data. There are widespread accusations in the media of voluntary, proactive and secret cooperation of telecom and internet providers with the NSA and/or other national intelligence or security agencies. It is difficult, if not impossible, to confirm these accusations based on hard evidence, since such evidence could not be identified or accessed by the authors of this study.

Representatives or spokespersons of all organisations that were publicly accused of being suspicious of having secretly cooperated with the NSA in particular, have strongly rejected these incriminations. Google's CEO Eric Schmidt rebutted in an interview to the US newspaper "ABC News" an accusation of Julian Assange, founder of the WikiLeaks platform, that Google is a "*privatized version of the NSA*", saying that "*Google never collaborated with NSA and in fact, we've fought very hard against what they did and since what the NSA did which we do not like, we have taken all of our data, all of our exchanges, and we fully encrypted them so no one can get them, especially the government*."[394]

Nine companies (AOL, Apple, DropBox, Facebook, Google, LinkedIn, MicroSoft, Twitter, Yahoo) among those who were accused of having secretly cooperated with the NSA have published a letter to the US Senate[395] in which they demand a Global Government Surveillance Reform.

The US Department of Justice has issued a statement on 27 January 2014[396] announcing new reporting methods for national security orders, which are expected to "*allow more detailed disclosures about the number of national security orders and requests issued to communications providers, and the number of customer accounts targeted under those orders and requests including the underlying legal authorities*". This statement implicitly admits that corresponding requests of information disclosure to communication providers have been, or are being issued. This leads to the conclusion that **although the cooperation between large communication and service providers and the NSA may not have occurred on a voluntary basis, data has been made accessible by these companies to the NSA on the basis of lawful requests, or under the threat of being fined.**

The "Washington Post" reports, for instance, that "*The U.S. government threatened to fine Yahoo US$250,000 a day in 2008 if it failed to comply with a broad demand to hand over user communications – a request the company believed was unconstitutional – according to court documents unsealed* [early

[394] http://abcnews.go.com/Business/googles-eric-schmidt-calls-julian-assange-paranoid-tim/story?id=25679642

[395] https://www.reformgovernmentsurveillance.com/

[396] http://www.justice.gov/opa/pr/joint-statement-attorney-general-eric-holder-and-director-national-intelligence-james-clapper

September 2014]"[397]. It is also known that a large number of requests for information disclosure have been and are being issued by the US Foreign Intelligence Surveillance Court (FISA), which has been criticized for its lack of oversight and public control mechanisms and which, according to information published by the Electronic Privacy Information Center[398], has only rejected 11 out of 33949 warrant requests for electronic surveillance in the period from 1970 to 2012.

4.2 Question 31

What hacking programmes and other counter-intelligence measures have been put in place by governmental agencies to avoid "going black" due to growing use of sophisticated encryption over the Internet?
This includes for instance NSA TAO and other similar programmes in the world.

The NSA elite cryptanalytic unit, the Office of Cryptanalysis and Exploitation Services (S31), a subdivision the Office of Tailored Access Operations (TAO), is composed of more than 1,500 cryptanalysts, mathematicians, scientists, engineers, and computer technicians who, according to information published on foreignpolicy.com[399] "*have had a remarkably large number of code-breaking successes against foreign targets*", which, however, "*were largely dependent on clandestine intelligence activities […] and not the more traditional cryptanalytic attacks on encrypted messages*". The same article states that NSA's cryptanalytic unit was largely unsuccessful trying to "*crack the encryption protection used by the anonymizer service Tor*", or files encrypted with PGP[400]. However, the latest developments concerning operation "Onymous"[401] in which Law Enforcement Agencies have managed to break security mechanisms of the TOR network in a way that even the proper developers of TOR were unable to identify[402], suggest that the cryptanalytic units have made some progress in this respect. This suspicion has been confirmed just recently in a research paper published by American and Italian computer scientists, who were able to "***reveal the actual sources of anonymous traffic*** *with 100% accuracy for the in-lab tests, and achieved an overall accuracy of about 81.4% for the real-world experiments*"[403]

The difficulties encountered for decrypting these protections by brute force attacks (despite the huge supercomputing capacity available to the named unit[404]) **has led the NSA to resort to clandestine techniques, the most productive of which is "*to covertly hack into targeted computers*"**, so as to be able to access and copy information before it is encrypted. For this purpose, the NSA employs an extensive set of tools[405] to covertly infiltrate computer networks and set up backdoors that allow them to take control over these networks.

Mikko Hypponen, a distinct computer security expert and Chief Research Officer at F-Secure, a Finish computer security firm, states that "[…] *some U.S.-based defence contractors have more than a hundred open positions for people with Top Secret/SCI clearance to create exploits. Some of these positions specifically mention the need to create offensive exploits targeting iPhones, iPads, and Android devices*"[406].

---

[397] http://www.washingtonpost.com/business/technology/us-threatened-massive-fine-to-force-yahoo-to-release-data/2014/09/11/38a7f69e-39e8-11e4-9c9f-ebb47272e40e_story.html
[398] http://epic.org/privacy/wiretap/stats/fisa_stats.html
[399] http://www.foreignpolicy.com/articles/2013/10/15/the_nsa_s_new_codebreakers
[400] http://en.wikipedia.org/wiki/Pretty_Good_Privacy
[401] http://en.wikipedia.org/wiki/Operation_Onymous
[402] https://blog.torproject.org/blog/thoughts-and-concerns-about-operation-onymous
[403] https://mice.cs.columbia.edu/getTechreport.php?techreportID=1545
[404] According to the article referenced in footnote 399, "*the NSA* [today] *spends over US$247 million a year to buy and maintain its state-of-the-art supercomputer systems just for cryptanalytic use*"
[405] https://www.eff.org/files/2014/01/06/20131230-appelbaum-nsa_ant_catalog.pdf
[406] CH@NGE 19 Key Essays on How Internet Is Changing Our Lives, Mikko Hypponen, Cyber Attacks, bbvaopenmind.com

According to a report on the NSA tool Treasuremap[407], the Deutsche Telekom (along with another telecom provider), which in Germany alone provides 60 million customers with mobile, Internet and fixed lines, are marked on a network diagram showing "*SIGINT collection points*", i.e. an access point for cyber spies in their network.

The "TS/SI/REL TO USA, FVEY" classified presentation, published through WikiLeaks, indicates that there exist 13 covered servers in unwitting data centres in 16 countries worldwide:

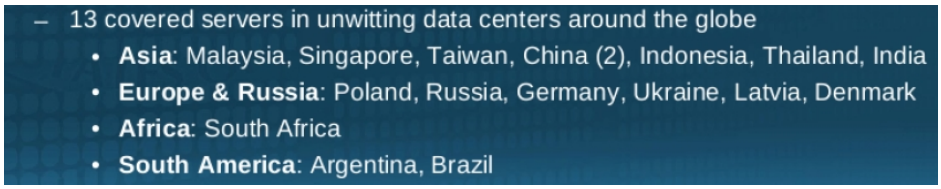

– 13 covered servers in unwitting data centers around the globe
• Asia: Malaysia, Singapore, Taiwan, China (2), Indonesia, Thailand, India
• Europe & Russia: Poland, Russia, Germany, Ukraine, Latvia, Denmark
• Africa: South Africa
• South America: Argentina, Brazil

Figure 28: Extract of slide #11 of Treasuremap presentation leaked by Snowden

But obviously not only the USA maintains such kind of cyber-espionage programs. Namely the UK, Canada Australia and New Zealand, as members of the FiveEyes (FVEY) intelligence alliance; China, Russia, Iran, Pakistan, but also EC member countries like Germany, France and Italy, among others, have been reported to run government-sponsored cyber-espionage units or programs. A recently leaked document on the customers[408] of a cyber-espionage application, distributed by the German company FinFisher, adds even more countries to this list. According to information published on WikiLeaks[409] "*FinFisher* (formerly part of the UK based Gamma Group International until late 2013) *is a German company that produces and sells computer intrusion systems, software exploits and remote monitoring systems that are capable of intercepting communications and data from OS X, Windows and Linux computers as well as Android, iOS, BlackBerry, Symbian and Windows Mobile devices*", which, according to a study of Citizen Lab[410], are being used in 36 countries worldwide, including the US and various EU member states.

An intelligence report from the American cyber-security firm Mandiant[411] claims that in China are operating "*more than 20 Advanced Persistent Threat (APT) groups*", one of which – codenamed APT1 – is being analysed in this report and "*is believed to be the 2nd Bureau of the People's Liberation Army (PLA) General Staff Department's (GSD) 3rd Department, which is most commonly known by its Military Unit Cover Designator (MUCD) as Unit 61398*." China insisted the reports were untrue and that they were not supporting any cyber-attacks.

CrowdStrike, another US security company announces on its web[412] that they "[…] *publicly released a report on a* [second] *group* […] *that conducts operations from Shanghai, China, likely on behalf of the Chinese People's Liberation Army (PLA) 3rd Department 12th Bureau Unit 61486.* [This group] *is a determined adversary group, conducting intelligence-gathering operations targeting the Government, Defense, Research, and Technology sectors in the United States, with specific targeting of space, aerospace, and communications*."

The cyber-attack on Estonia in 2007, in which a number of websites of Estonian organizations, including the parliament, banks, ministries, newspapers and broadcasters were subject to a Distributed Denial of Service (DDoS) attack, was considered as one of the most sophisticated attacks carried out up to that date. Many observers and politicians accused the Russian Government to be the driving force behind the attack, but experts have never been able to proof implication of the Russian Government in the case.

[407] https://firstlook.org/theintercept/document/2014/09/14/treasure-map-presentation/
[408] https://wikileaks.org/spyfiles4/customers.html
[409] https://wikileaks.org/spyfiles4/
[410] https://citizenlab.org/2013/04/for-their-eyes-only-2/
[411] http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf
[412] http://resources.crowdstrike.com/putterpanda/

A report by The Guardian on the Russian secret service monitoring 'all communications' at Winter Olympics in Sochi, cites Ron Deibert, a professor at the University of Toronto and director of Citizen Lab, saying that "*the scope and scale of Russian surveillance are similar to the disclosures about the US program* […]"[413].

An article written by Jennifer Hesterman[414], a retired USAF colonel, claims that, despite having a shorter hacking history, "*in May 2013, Iranian hackers were able to access the computer networks running U.S. energy companies, possibly giving Iran the means to sabotage power plants" and that "Iran is believed to be actively attacking the U.S. power grid and military*". Hesterman does however not back this assertion with references or proof of evidence.

The part of the Israeli Defence Forces (IDF) which is responsible for SIGINT collection is called Unit 8200. This unit only recently got a lot of media coverage due to a letter that was published by 43 of its soldiers and in which they criticize intelligence gathering practices that "*harm innocents and serve for political persecution and sowing discord in Palestinian society*."[415] Despite the fact that little is known about this IDF special unit, reports on its recruitment practices[416] in which they design hacking contests "*in which students were required to build a bot (an automated program) designed to attack "enemy" servers, while defending their own servers from attack*", lead to the conclusion that the goals, methods and techniques applied by the IDF for SIGINT collection do not substantially differ from those employed by other intelligence agencies.

According to an article in the German magazine "Der Spiegel"[417], the German intelligence service BND is running a SIGINT collection facility in Bad Aibling which has been taken over from the US Army in 2004. The article states that the BND believes Bad Aibling is one of two Signals Intelligence Activity Designators (SIGADs) that are mentioned in documents leaked by Edward Snowden and "*through which the controversial US intelligence agency gathered about 500 million pieces of metadata in December 2012 alone*". The article also suggests that there exist tight links between the NSA and the BND, which include the exchange of SIGINT data and access to surveillance applications. There is, however, no notion of the BND employing hacking techniques or units to gain intelligence information.

The question remains on how such secret cyber-espionage and hacking programs can be kept alive without being subject to leaks. The Snowden case – among other whistle-blower cases of lesser media coverage – shows that such leaks cannot be completely avoided. But the different national intelligence agencies do of course establish various means for preventing such leaks to happen.

To start with, the NSA requires all employees to sign a Classified Information Nondisclosure Agreement[418] in which the signer confirms that "*I have been advised that any unauthorized disclosure of classified information by me may constitute a violation, or violations, of United States criminal laws* […]". This section is basically the one that US law enforcement refers to when arguing that Snowden has breached "his oath" and that he cannot be considered a whistle-blower under the Military Whistle-blower Protection Act or other whistle-blower protections, which are also referred to in the mentioned NDA.

Although no information is available on the respective practices in other national security agencies, it can be safely assumed that their employees are required to sign similar NDAs.

[413] http://www.theguardian.com/world/2013/oct/06/russia-monitor-communications-sochi-winter-olympics
[414] http://onlinedigitalpublishing.com/display_article.php?id=1610251
[415] http://www.haaretz.com/news/diplomacy-defense/.premium-1.615811
[416] http://www.timesofisrael.com/for-hack-contest-winners-a-ticket-into-unit-8200/#!
[417] http://www.spiegel.de/international/world/german-intelligence-sends-massive-amounts-of-data-to-the-nsa-a-914821.html
[418] http://www.archives.gov/isoo/security-forms/sf312.pdf

Secondly, security agencies, like any other military or national defence organization, are organized in a strictly hierarchical structure in which the information is compartmented in such a way that only very few unique persons (if at all) do have a total oversight of all activities. This means that even if a specialist is working on a critical issue (e.g. related to surveillance) she may not be able to draw the connection to related actions and get the "big picture" of what she is contributing to ("[…] *in many of our foreign partners' capitals few senior Officials outside of their defence-intelligence apparatuses are witting to any SIGINT connection to the U.S./NSA*")[419]. This makes it much more difficult for an individual to gather sufficient information or insight in complex programs and therefore reduces the risk of someone becoming suspicious and/or being able to leak relevant and connected information.

Another influencing factor that prevents security relevant information to be leaked is the public opinion, or stance, in some countries. While the public opinion in many European countries and Canada tends toward defending the whistleblowing of Edward Snowden, for instance, the US American society's opinion is split. According to different polls, there is hardly any difference between the percentage of people who consider Snowden a "hero" or a "traitor"[420],[421],[422]. The major differences exist between groups of different age, with the younger generation defending the opinion that Snowden has served the public interest, and the older generation rather seeing him as a traitor. This may be related to different perceptions of societal values and principles in different nations. The comment of the Director of the NSA on the priority of civil rights or security on NSA's website potentially helps to understand this argument[423].

4.3 Question 32

What can be concluded regarding the efficiency/inefficiency of these programmes taking into account information publicly available, and the state of the art of science and technology?

The strategies applied by security agencies for gather SIGINT data can be classified in 4 categories:

- Network intrusion (hacking)
- Infection of devices with secret backdoors (troyans, keylogger, etc.)
- Physical modification of hardware components
- Weakening, or breaking, publicly trusted encryption standards

Network Intrusion can reap a wealth of information, but more importantly, gives the attacker the ability to re-configure and "steer" the network from inside. A screenshot of an email conversation[424] from the pool of Snowden documents, published by the Electronic Frontiers Foundation, documents in a bullet list the advantages of hacking large network routers ("*such as CISCOs / Junipers / Huaweis*"), which include the possibility to

- add credentials, allowing a third party to log in anytime
- add/change routing rules (e.g. detour traffic to servers controlled by third parties)
- setup a packet capture capability (e.g. for "sniffing" credentials)

[419] https://www.eff.org/files/2014/04/09/20140313-intercept-nsa_cooperative_third_party_relationships.pdf
[420] http://www.angusreidglobal.com/wp-content/uploads/2013/10/2013.10.30-Snowden-Leaks.pdf
[421] http://www.people-press.org/2013/06/17/public-split-over-impact-of-nsa-leak-but-most-want-snowden-prosecuted/
[422] http://www.pewresearch.org/fact-tank/2014/01/22/most-young-americans-say-snowden-has-served-the-public-interest/
[423] https://www.nsa.gov/about/values/core_values.shtml (ref. Q5)
[424] https://www.eff.org/files/2014/04/09/20140312-intercept-five_eyes_hacking_large_routers.pdf

- weaken any VPN encryption capabilities on the router, forcing it to create tunnels that can easily be decrypted
- install a modified version of the Operating System with whatever pre-built-in functionality

A practical example of how far such possibilities can go is the case of Stellar PCS, a German Satellite Communications Provider that provides internet access to remote portions of the globe via satellite. When engineers of Stellar PCS were confronted with the contents of a leaked document on GCHQ's Treasuremap[425] application that showed the level of infiltration of the GCHQ into their network (including access passwords of other providers who rented Stellar PCS's infrastructure), the system administrator of Stellar PCS said that with this access level the GCHQ would be able to change links, geo-locate users, or "[…] *shut down the internet in entire African countries that are provided access via our satellite connections*"[426].

Infection of devices with secret backdoors is a technique that allows eavesdropping information without the knowledge of the owner or user of an infected system**.** Such backdoors can be established by purposefully modifying software or by taking advantage of (unintended) software flaws[427].

Documents released by Snowden reveal surveillance powers that go beyond root access and into the hardware of different kinds of network devices, servers and handheld devices. The catalogue of NSA's Tailored Access Operations (TAO) group, listing a number of exploits, can be accessed on the EFF website[428].

Dr. Matteo Bonfanti, Research Fellow in International Law and Security at Scuola Superiore Sant'Anna Pisa (Italy) and expert on Ethics, considers that "*installing "backdoors" in the hardware and software systems jeopardizes not only users' trust in technologies, but also the market in general*".

Physical modification of hardware components is a rather sophisticated technique that is difficult, if not impossible to detect without advanced digital forensic knowledge and tools. Although not directly ascribed by evidence to being used by security agencies today, a technique called "transistor doping"[429] is able to alter the correct behaviour of integrated circuits by modifying the crystalline structure of transistors.

Researchers of the University of Massachusetts have shown that by manipulating the Random Number Generator (RNG) that provides "*the starting random numbers with which to create encryption keys*", they were able to change the behaviour of the chip "*so that one particular number became a constant instead of a variable*[430]*. This weakens any encryption that comes from keys created by that system*" and consequently facilitates decryption by third parties that don't possess the full decryption key. According to the researchers, such a hardware modification would not be detected by the chips' self-test and be virtually impossible to be detected by users.

An article in IEEE Spectrum[431] from 2008 reports that the US Department of Defense (DOD) launched the "Trust in Integrated Circuits program", with the aim of verifying the integrity of the integrated circuits (IC) that form part of military systems. The main concern that justifies this program is the DODs fear that chips which "*have been purposely fabricated with a hidden "backdoor" inside*" could be integrated in their military equipment and give enemy forces the possibility to remotely control or sabotage them (e.g. by activating so called "kill-switches").

[425] https://firstlook.org/theintercept/document/2014/09/14/treasure-map-presentation/
[426] https://firstlook.org/theintercept/2014/09/14/nsa-stellar/
[427] See also Question 8 and Question 10
[428] https://www.eff.org/files/2014/01/06/20131230-appelbaum-nsa_ant_catalog.pdf
[429] http://en.wikipedia.org/wiki/Doping_(semiconductor)
[430] http://www.techrepublic.com/blog/it-security/researchers-create-nearly-undetectable-hardware-backdoor/
[431] http://spectrum.ieee.org/semiconductors/design/the-hunt-for-the-kill-switch

There are reported cases[432] of devices from Asian computer manufacturers being banned from networks of security agencies, because of "*concerns about security threats posed by "malicious circuits" and insecure firmware in chips produced in China by companies with close government ties*".

A common practice, which has purportedly been used to infiltrate the Stuxnet virus in Iranian nuclear facilities, is that of modifying the firmware of USB flash drives (pen drives). This technique consists in reprogramming the USB controller chip so that it can infect the host system with malicious code[433]. There is little to no protection available (except for clean-room or white-room separation of systems from external devices) for preventing attacks that employ this infection technique, since it does not get detected by malware or antivirus scanners and since specific USB devices cannot (yet) be blocked by firewalls. It may not even be possible to eliminate injected vulnerabilities by reinstalling the operating software, if the attack code modifies the BIOS of the affected system.

Weakening, or breaking, publicly trusted encryption standards is a strategy that requires influencing the standardisation bodies or getting in secret alliance with the vendors that implement these standards and much has been written about the effort of security agencies in gaining access to the keys that would allow deciphering information that is encrypted by different encryption technologies[434]. This battle goes back to the early 90s, when Phil Zimmerman, the creator of PGP, a public domain cryptology application, was subject of a criminal investigation by the US Government, accusing him of exporting munition without license (Cryptosystems using keys larger than 40 bits were then considered munitions within the definition of the US export regulations[435])

More recent information published by Reuters blames the NSA of having paid US$10 million to the RSA, an American computer and network security company, in a deal that "*promulgated a flawed formula for generating random numbers*" in RSA's BSafe software, "*that is used to enhance security in personal computers and many other products*"[436] This flaw would facilitate cracking encryptions that have been scrambled by the "Dual Elliptic Curve Deterministic Random Bit Generator" (Dual EC DRGB) algorithm used by BSafe. This publication led a number of high-level speakers "*to withdraw from the 2014 RSA Conference in San Francisco, which attracts some 25,000 attendees each year*"[437].

The RSA has categorically denied the allegation of having "*entered into a "secret contract" with the NSA to incorporate a known flawed random number generator into its BSAFE encryption libraries*". The NSA did not comment on the accusations.

However, other information published by New York Times, the Guardian and ProPublica does sustain and even extend the accusation of the NSA secretly defeating internet privacy and security[438], revealing that

- a 10-year NSA program against encryption technologies made a breakthrough in 2010 which made "*vast amounts*" of data collected through internet cable taps newly "*exploitable*".
- the NSA spends 250 million US dollar a year on a program which, among other goals, works with technology companies to "*covertly influence*" their product designs.
- a GCHQ team has been working to develop ways into encrypted traffic on the "big four" service providers, named as Hotmail, Google, Yahoo and Facebook

[432] http://www.afr.com/p/technology/spy_agencies_ban_lenovo_pcs_on_security_HVgcKTHp4bIA4ulCPqC7SL
[433] https://srlabs.de/badusb/
[434] See also Section 2
[435] http://en.wikipedia.org/wiki/Pretty_Good_Privacy
[436] http://www.reuters.com/article/2013/12/20/us-usa-security-rsa-idUSBRE9BJ1C220131220
[437] https://www.eff.org/deeplinks/2014/01/after-nsa-backdoors-security-experts-leave-rsa-conference-they-can-trust
[438] http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security

A statement from the Office of the Director of National Intelligence on the reported information, which is published on ProPublica says that "[…] *the fact that NSA's mission includes deciphering enciphered communications is not a secret, and is not news. Indeed,* ***NSA's public website states that its mission includes leading "the U.S. Government in cryptology ... in order to gain a decision advantage for the Nation and our allies***"." It also adds that the information disclosed in the referred media "*reveal specific and classified details about how we conduct this critical intelligence activity. Anything that* [these] *disclosures add to the ongoing public debate is outweighed by the road map they give to our adversaries about the specific techniques we are using to try to intercept their communications in our attempts to keep America and our allies safe and to provide our leaders with the information they need* […]"[439].

In an effort to maintain the mentioned decision advantage by means of technological advantage, the NSA, according to the Washington Post, is performing a 79.7 million US dollar research program titled "Penetrating Hard Targets"[440] with the objective to build a quantum computer that is able to break the majority of the encryption standards on the Internet. It must however be pointed out that quantum computing today is still rather a theoretical concept than a ready-to-implement engineering solution and that "*experts predict that it would take at least five more years to attain the kind of quantum computers that the NSA wants*".

The documented facts and information lead to the conclusion that **Governmental agencies do in fact possess capabilities to break virtually any system protection and to infiltrate systems and networks at their convenience by applying (a mix of) state of the art technology. Even though it is possible to hinder such unauthorized intrusions by applying different protection mechanisms, there is no means for guaranteeing total immunity against such attacks.**

The highest level of protection is provided by end-to-end encryption – as long as it can be excluded that potential attackers know the decryption key, or part of it, and as long as it can be excluded that attackers can access information before it gets encrypted by the system. Such an encryption can however only protect data content. But it is all but impossible to prevent information be derived from the observation of – generally unencrypted and often publicly accessible - meta-data. As recent studies have shown, this kind of meta-data analysis can "*yield vast amounts of information*" that reveal "*a wealth of detail about family, political, professional, religious and sexual associations*"[441].

4.4 Question 33

> ***Are all the statements made by the press and whistle-blowers systematically technically coherent and consistent?***
> ***What is the technical credibility and global coherence of the technical statements made publicly in the press following Edward Snowden's revelations?***

Despite the rebuttal of many accusations, particularly those related to collaborations between commercial internet companies and national security agencies, the authors of this report were not able to identify technical rebuttals of the revealed NSA documents, neither through revision of literature, nor through the statements of technical experts on the respective subject matter. Although this absence of technical refutations cannot be equated to being a validated proof of credibility or technical coherence of Snowden's revelations, it leads to the conclusion that **the technical feasibility of the tools and practices applied by the NSA (and other national security agencies) is not disputed by any of the relevant technical communities.**

[439] http://www.propublica.org/article/the-nsas-secret-campaign-to-crack-undermine-internet-encryption
[440] http://www.washingtonpost.com/world/national-security/nsa-seeks-to-build-quantum-computer-that-could-crack-most-types-of-encryption/2014/01/02/8fff297e-7195-11e3-8def-a33011492df2_story.html?hpid=z1
[441] http://news.stanford.edu/news/2014/march/nsa-phone-surveillance-031214.html?_ga=1.80287443.511274451.1412847227

The statements that can be found in some of the original leaked documents directly contribute to confirm the global coherence and consistency of the published accusations. So does a slide (see Figure 29) from the fund of Snowden documents that presents the "New Collection Posture" of the NSA with the six bullet points "*sniff it all, know it all, collect it all, process it all, exploit it all and partner it all*"[442]. This can only be interpreted as the NSA having "*explicitly embraced* [General Keith B.] *Alexander's*[443] *motto of omniscience as its core purpose*"[444].

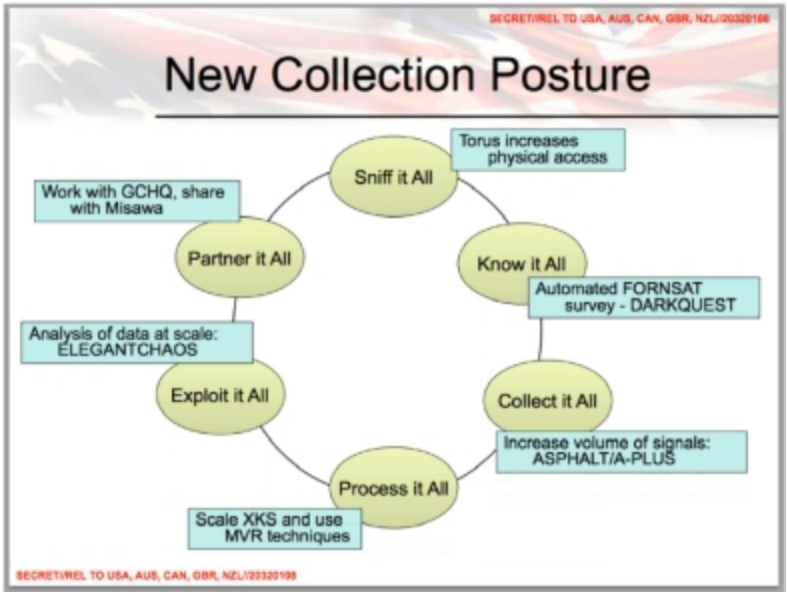

New Collection Posture
Sniff it All
Torus increases physical access
Work with GCHQ, share with Misawa
Partner it All
Know it All
Automated FORNSAT survey - DARKQUEST
Analysis of data at scale: ELEGANTCHAOS
Exploit it All
Collect it All
Increase volume of signals: ASPHALT/A-PLUS
Process it All
Scale XKS and use MVR techniques

Figure 29: Leaked slide of a presentation to the 2011 annual conference of the Five Eyes alliance on NSA's SIGINT gathering strategy

Furthermore, the reports on indiscriminate collection of private information on the web, which have neither been denied nor admitted by the NSA, do substantiate the validity of the allegations published by the media in relation with the Snowden documents:

The New York Times reports[445], based on a document obtained from Edward Snowden, that The National Security Agency is harvesting "*millions of images* [of people] *per day*" — including about 55,000 "*facial recognition quality images*", from communications that it intercepts through its global surveillance operations. The article states that "*the agency's use of facial recognition technology goes far beyond one program previously reported by The Guardian, which disclosed that the NSA and GCHQ have jointly intercepted webcam images, including sexually explicit material, from Yahoo users"*. This practice does not seem to be unlawful in the US, since their privacy laws provide no express protections for facial recognition data.

The Washington Post adds further details to this practice in an article[446] from July 2014 in which it states that "*nine of 10 account holders found in a large cache of intercepted conversations* […] *were not the intended surveillance targets but were caught in a net the agency had cast for somebody else*" and that much of this intercepted information has "*a startlingly intimate, even voyeuristic quality.*" The same article nevertheless also confirms that "*months of tracking communications across more than 50 alias accounts* […]

[442] https://edwardsnowden.com/wp-content/uploads/2014/05/newcollectionposture.pdf
[443] Four-star General Keith Brian Alexander was the Director of NSA from 2005 to 2013. He also held the position of Chief of the Central Security Service (since 2005) and Commander of the United States Cyber Command (since 2010)
[444] Greenwald, G.: No Place to Hide: Edward Snowden, the NSA, and the U. S. Surveillance State, Metropolitan Books, New York (2014)
[445] http://www.nytimes.com/2014/06/01/us/nsa-collecting-millions-of-faces-from-web-images.html?_r=2
[446] http://www.washingtonpost.com/world/national-security/in-nsa-intercepted-data-those-not-targeted-far-outnumber-the-foreigners-who-are/2014/07/05/8139adf8-045a-11e4-8572-4b1b969b6322_story.html

led directly to the 2011 capture in Abbottabad of Muhammad Tahir Shahzad, a Pakistan-based bomb builder, and Umar Patek, a suspect in a 2002 terrorist bombing on the Indonesian island of Bali".

A top secret NSA document of July 2010, accessible on EFFs website[447], lists the Nations and Organizations that were authorized surveillance targets at the time. Out of the 195 independent states listed on nationsonline.org[448], the NSA list considers 194 as legitimate surveillance targets. Or, in other words, the only nation the USA is excluding from this list is the USA itself. Nevertheless, this does not imply that no SIGINT data is or can be collected by the NSA on US citizens. "The Guardian" reports on secret (FISA) court orders that open way for domestic surveillance also in the US and with US citizens being targeted.[449]

As already mentioned in Question 32, the NSA has launched a 79.7 million US dollar research program that investigates in quantum computing technology, which, despite the fact that it is currently still largely in its conceptual state, could become a strategic advantage with respect to decryption capabilities in the future. Although no further information on similar investments of the NSA in other research or technology areas is available, it can be safely assumed that the agency (and their international counterparts) is thoroughly screening and investigating research and technology trends that could provide a strategic advantage in the future for obtaining SIGINT data.

Given the fact that national security counts with an enormous budget - according to an interactive graphic published by the Washington Post, 4.3 billion US dollar are assigned in the US only for conducting cyber operations[450] - it is more than likely that such research endeavours will allow to reap their benefits ahead of other civil sectors. The technologies that are most prone to be subject of such an investment are clearly those that would

- increase computing power by order of magnitude for decryption purposes (e.g. quantum computing, graphene chips)
- allow to modify integrated circuits at their hardware level (e.g. transistor doping[451])
- allow to (semantically) analyse large amounts of data (e.g. big data analysis, facial recognition)
- allow to identify potential (zero-day) exploits that can be used to insert vulnerabilities in computer systems and networks

The Quadrennial Intelligence Community (IC) Review Report of 2009[452] identifies "*three concepts ("safe bets") that appear critical to the success of the IC across a wide range of possible futures*" and three "*Strategic Hedges, concepts deemed highly relevant in some possible futures but less relevant in others*".

These six concepts are:

- Safe bets
    - sensing and learning environment for humans and intelligent machines to analyse "exabytes" of data in near-real time to generate and test hypotheses, autonomously process and evaluate insights to cue collection, and self-update/self-correct.
    - more customized tools, products and services to an expanding set of customers with different styles and end uses.

[447] https://www.eff.org/files/2014/06/30/list_of_governments_and_organizations_authorized_for_surveillance.pdf
[448] http://www.nationsonline.org/oneworld/states.htm
[449] http://www.theguardian.com/world/2013/jun/20/fisa-court-nsa-without-warrant
[450] http://www.washingtonpost.com/wp-srv/special/national/black-budget/
[451] https://www.cl.cam.ac.uk/~sps32/ches2012-backdoor.pdf

[452] https://www.eff.org/files/2014/09/22/20140905-intercept-quadrennial_intelligence_community_review_final_report.pdf

  - expeditionary intelligence capacity to deploy small teams to hostile (physical or virtual) environments with minimal infrastructure, while readily drawing on more diverse and instantly collaborative analytic expertise to guide operations and directly leverage the insights gained.

- Strategic Hedges
  - aggressively employing a mix of overt means, clandestine penetration, and counterintelligence tactics to address severe U.S. technological erosion vis-à-vis near-peer competitors and global corporations.
  - confronting environments dominated by non-state actors, requiring unconventional human collection methods with more flexible sets of analytic partners to track highly empowered, cyber-immersed individuals and groups.
  - compensate for the possible loss of access to reliable financial and economic data (at the global, national, and sector levels) by penetrating corporations, foreign finance ministries, central banks, and market participants to create an "economic operating picture."

4.5 Question 34

How efficient are all the security solutions available on the market to fight mass surveillance by hacking programmes of national security agencies? How efficient is the use of "end to end" encryption to fight back the possible cooperation of mass surveillance organisations with telecom operators, Internet service providers, providers of cloud, web, email, social network and voip services?

The **security solutions on the market** available for users to protect themselves against any type of surveillance fall basically into five categories:

- antivirus programs
- firewalls
- VPNs
- encryption tools
- anonymizer services and tools

They all **provide a certain level of security** for routine and conventional threats, impeding unwanted privacy invasion, or infection with vulnerabilities by third parties, **but even if applied in combination they can't guarantee total immunity against complex attacks**.

Antivirus programs do provide a good level of protection for known malware, viruses, trojans and even malicious URLs, spam, or rootkits. They do apply different strategies for detecting (and eventually disinfecting) malicious software, which are based on the identification of signatures (or patterns), or on heuristic methods.

A **firewall** is either a software application or a hardware appliance that can block in- and outgoing network traffic on a device, based on defined rules and depending on the communication ports and/or protocols used.

A **Virtual Private Network (VPN)** is, as its name suggests, a private (virtual) network that makes use of public network structures, requiring authenticated access and employing different security and encryption techniques, in order to guarantee the privacy of data exchanged between two endpoints.

Encryption tools are software applications that encrypt and decrypt data or communication channels, applying different algorithms and cipher suites.

Anonymizer services and tools are proxies that provide anonymity and privacy for users when accessing servers in the internet. This is achieved by obscuring the client's IP address and the "path" that was taken for accessing a server. This makes it impossible for third parties to gather information about which servers a user is accessing and for the accessed server to gather the IP address of the client accessing it.

All of these security solutions have their particular weaknesses that could allow attackers to break their protection.

Antivirus programs may fail when it comes to zero-day-attacks. If such a zero-day-attack makes use of new techniques (or signatures/patterns) that are not covered by the knowledge base of the antivirus application, it may not be detected.

Firewalls can be bypassed by applying different types of attack (e.g. MITM, DNS spoofing, DDoS, Rootkits, etc.). VPNs also can be compromised in various ways[453] and are subject to attacks that "*take advantage of misconfigurations and poorly managed implementations*"[454].

Encryption, which at first sight may appear to be one of the strongest means to protect form privacy violations, is only as good as the algorithms and cipher sets used. Despite the fact that encryptions employing keys of 128bits or longer are theoretically impossible to decipher with today's available computing power, a number of recent reports suggest that some security agencies – first and foremost the NSA and GCHQ – have made significant progress in cracking certain types of encryption so that "*vast amounts of encrypted internet data which have up till now been discarded are now exploitable*"[455].

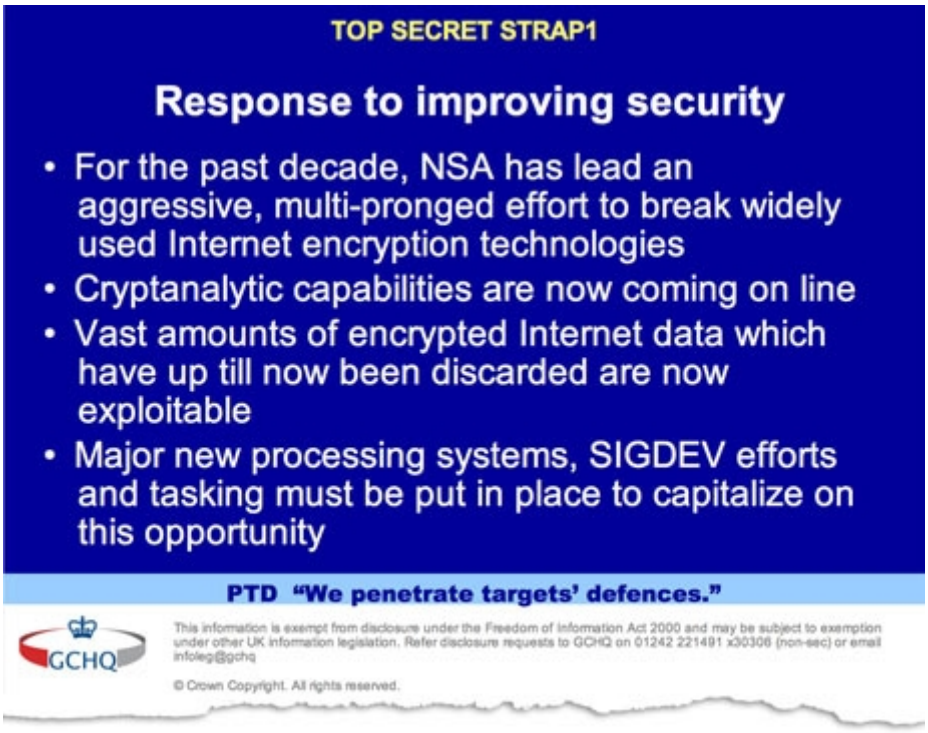

TOP SECRET STRAP1
Response to improving security
• For the past decade, NSA has lead an aggressive, multi-pronged effort to break widely used Internet encryption technologies
• Cryptanalytic capabilities are now coming on line
• Vast amounts of encrypted Internet data which have up till now been discarded are now exploitable
• Major new processing systems, SIGDEV efforts and tasking must be put in place to capitalize on this opportunity
PTD "We penetrate targets' defences."
GCHQ
This information is exempt from disclosure under the Freedom of Information Act 2000 and may be subject to exemption under other UK information legislation. Refer disclosure requests to GCHQ on 01242 221491 x30306 (non-sec) or email infoleg@gchq
© Crown Copyright. All rights reserved.

Figure 30: Leaked slide on GCHQs strategy on breaking encryption technology

[453] http://www.nta-monitor.com/files/whitepapers/VPN-Flaws-Whitepaper.pdf
[454] https://www.ncp-e.com/fileadmin/pdf/techpapers/NCP-Attack-Vectors-WP.pdf
[455] http://www.theguardian.com/world/2013/sep/05/nsa-gchq-encryption-codes-security

And finally, anonymizing services, namely the TOR network, one of the last technology resorts that were considered to effectively provide privacy, was only recently compromised in a joint operation of international law enforcement agencies – labelled operation "Onymous" - that targeted illegal websites which are only accessible through the TOR network[456]. Other reports are claiming that the Russian Government has posted a tender[457] in which it seeks for parties willing to study the possibility of obtaining technical information about the TOR anonymous network. This adds to the fact that anonymizing services require the data to be transmitted from the user's machine to the service. In these hops users' information is in transit so it may be intercepted by other means, irrespective of the intended effect of the service.

The ability of a user to effectively own and administrate private data furthermore conflicts with the use of cloud services. In this case the data will be entirely submitted to the policies and mechanisms of the cloud provider. Some providers provide backend services for encrypting data (which could also be prone to weak encryption algorithms or keys), while others leave encryption to the responsibility of the user (which reduces the risk of the provider being able to decrypt this data). Also a variety of metadata can be acquired from the user activity in the Cloud, since data manipulation (for example, growing in size), sharing, and uploading/downloading can be analysed to determine certain behaviour patterns.

It can therefore be concluded that the application of **a combination of the above described security and privacy solutions does provide an advanced level of protection** from mass surveillance practices **and** that **end-to-end encryption of communication channels does constitute a practically unbreakable security mechanism**. This does, however, not imply that total anonymity can be achieved by applying these security solutions and strategies. As has been documented, security and intelligence agencies (and potentially also some criminal organisations) do have the knowledge, expertise and technology for breaking even apparently sophisticated security and privacy protections (see also Question 20).

4.6 Question 35

What would be the advantages and disadvantages for national and European public administrations to continue to trust and massively use operating systems and applications edited outside of the EU by commercial organisations that are suspected of implementing legal backdoors upon the legal requests of their international security authorities?
What realistic measures could be taken to deter security threats?
Should the EU promote dissemination of new generation of secure open source operating system(s) as a matter of European policy?

The major advantages of using commercial software and equipment are related to the fact that they are thoroughly tested for their operation, are being continuously updated and that their supplier is liable for potential consequences of malfunctions, errors and bugs. In the case of some OS and application (suites), their vast deployment has led to a "de facto standard", which is even further leveraged through agreements with hardware manufacturers. Millions of users, including large parts of public administration worldwide, are using Microsoft Windows and Microsoft Office in their daily work. Interoperability issues with alternative OSs and applications, for instance in the Open Source software market, along with the costs of migration are the primary reasons for many organisations to stick with the status quo.

[456] http://www.theguardian.com/technology/2014/nov/11/operation-onymous-flaws-tor
[457] http://zakupki.gov.ru/epz/order/notice/zkk44/view/common-info.html?regNumber=0373100088714000008

One of the biggest concerns when using commercial software and equipment is their "black-box" character (i.e. their internal functioning is unknown to the user) and the consequent security issues that may derive from it. Open Source software does eliminate this concern through the openness of its code, which allows for constant inspection and scrutiny by a large community of experts. This does, however, not guarantee that Open Source software is error free (or less prone to errors than commercial software). The "heartbleed" bug in the OpenSSL protocol implementation, which is deployed in millions of servers worldwide, had remained for years in the source code before it was detected and fixed[458]. This means that **while the level of confidence in adequate security levels of commercial software and equipment is based on trusting the provider, in the case of Open Source software and equipment this confidence can be built on a user's own inspection and scrutiny.**

The US Department of Homeland Security has published a catalogue of Open Source security software that has been developed in the Homeland Open Security Technology (HOST) project.[459]

An example of an administration having migrated their IT infrastructure to Open Source is the German City of Munich[460]. **The goal of this migration was clearly to achieve more autonomy from providers** and that lower IT-costs were merely a "*welcome side-effect*". Their experience show that such a migration "*requires careful analysis and planning, as well as a clearly defined goal"* and that a lot of testing was required before introducing new tools, so as to ensure that the operation of the administration did not get interrupted. They also state that the "*lack of acceptance and outright resistance* [of the involved staff] *can be more substantial obstacles to the deployment of a software solution than any technical problem*"

Some experts claim that it is impossible to guarantee security and privacy while using Windows SO, since the US administration's close ties to MS – as an American company – will always give them a competitive edge over other countries when it comes to exploiting vulnerabilities in the OS code. This does NOT imply that backdoors are built in purposefully, but that the perpetual pipeline of zero day vulnerabilities caused by (unintended) software flaws opens way for backdoors. The crucial question is therefore who has knowledge of such zero day vulnerabilities (and the potential to exploit them) for how long before they become public.

Initiatives like the Qube-OS project try to address this issue by developing Open Source Operating Systems that are designed to provide strong security, which in the case of Qube OS is achieved by applying a "**Security by Compartmentalization"** approach.

Another critique voiced by security and privacy advocates is that Europe does too little to encourage the development of security and privacy related tools and that the available funding instruments (i.e. H2020) insist in that such applications need to be cross-platform, which in itself is an "incongruent idea". The information exposed in this document makes very clear that **achievable levels of security and privacy are very much depending on the specific platforms and technologies used** and that **security and privacy issues must therefore be addressed in a platform-specific way**.

Among sensible recommendations for blocking unwanted intrusion in personal communications stands out the use of public-domain encryption that has to be compatible with other implementations. It is harder for an intruder to backdoor any tool that has to be compatible with every other vendor's than one that only has to be compatible with itself. Other **recommendable preferences are symmetric cryptography over public-key cryptography, and conventional discrete-log-based systems over elliptic-curve systems**, as the latter have constants which compromise randomness. Finally, it is possible to use a system that allows the private key changing the message that has been sent without altering the encrypted code.

[458] See also Question 7

[459] http://www.dhs.gov/sites/default/files/publications/csd-host-open-soruce-cybersecurity-catalog.pdf

[460] https://joinup.ec.europa.eu/elibrary/case/declaration-independence-limux-project-munich-0

Concerned individuals use in-home elaborated architectures in an attempt to **isolate data storage from network connection** plus additional properties like AES cryptography, operating systems launched from live CDs, and disposable mail services. At a wider lever, some **strategies to provide sufficient anonymity may entail complex global structures, such as a worldwide service with multiple intermediates and end-points strengthened with randomly introduced delays and dummy encrypted data that flows constantly between the intermediate servers and end-points**. This would provide some form of anonymity much in the same way as the TOR network. Other initiatives underway are based on the P2P protocol. For example, Bitmessage[461] uses strong authentication to prevent spoofing of the sender and aims at hiding non-content data too, while relying on decentralization and avoiding dependence in certification authorities, whose dispersion and credibility has been put at stake.

Speakers of the IETF[462] have presented the organization in favour of applying encryption to all web traffic, starting by making encryption mandatory for HTTP 2.0, a new version of the basic web protocol. Additionally, the IETF plans to strengthen existing algorithms behind encryption while removing weak algorithms, some of them promoted by the same agencies suspect of further breaking them. All of this bearing in mind that the IETF might be able to secure the pipes through which users' data travel, but that ultimately users must also be able to trust the parties where their data is stored: software, hardware and services such as routers, remote mail storage, or social media.

Networks exposing a simple layer for anonymous message exchange between applications do already exist. In the Invisible Internet Project (I2P)[463] "*all communication is end to end encrypted (in total there are four layers of encryption used when sending a message), and even the end points ("destinations") are cryptographic identifiers (essentially a pair of public keys)*".

One latent problem is that many of the most widely used cipher suites do not support perfect forward secrecy, meaning that if a server's long-term key is compromised then all traffic ever encrypted with it can be broken. **Perfect forward secrecy means that any recorded traffic from past sessions remains useless even if the long-term server keys are subsequently revealed. It seems essential that perfect forward secrecy becomes universally used.**

[461] https://bitmessage.org/wiki/Main_Page
[462] https://www.ietf.org/
[463] https://geti2p.net/en/

This document contains the Annex to the first part of the study on Mass surveillance, commissioned by STOA.

This Annex provides detailed information on the four subthemes defined in the initial invitation to tender for the study.

This is a publication of the
Directorate for Impact Assessment and European Added Value
Directorate-General for Parliamentary Research Services, European Parliament

PE 527.409
ISBN: 978-92-823-5537-4
DOI: 10.2861/566206
CAT: QA-05-14-156-EN-N